# JOURNAL OF APPLIED PHYSICS

# JOURNAL OF APPLIED PHYSICS

Vol. 20, No. 7 JULY, 1949


## *In This Issue*


| | Page |
|---|---|
| **Compressibility Effects in Two-Phase Flow** R. F. Tangren, C. H. Dodge, and H. S. Seifert | 637 |
| **Fluid Flow Patterns** Garrett Birkhoff and Thomas E. Caywood | 646 |
| **Dust—An Active Meteorological Factor in the Atmosphere of Northern Africa** M. G. El-Fandy | 660 |
| **Some Factors Affecting the Rate of Precipitation Hardening in Cu-Be Alloys** Paul A. Beck | 666 |
| **Semi-Conducting Properties in Oxide Cathodes** N. B. Hannay, D. MacNair, and Addison H. White | 669 |
| **Thermionic Emission from Sintered Cathode of Thoria and Tungsten Mixture** H. Y. Fan | 682 |
| **High Temperature Furnace for Electron Diffraction Studies of Thin Films** Eileen I. Alessandrini | 691 |
| **Method of Equivalent Linearization for Non-Linear Oscillatory Systems with Large Non-Linearity** Carl A. Ludeke | 694 |
| **Illuminating System of the Electron Microscope** James Hillier and S. G. Ellis | 700 |
| **Point-to-Plane Corona Onsets** William N. English and Leonard B. Loeb | 707 |
| **Magnetic Properties of Iron Compacts in Relation to Sintering Temperature** Robert Steinitz | 712 |
| **Vaporization of the Cathode in the Electric Arc** Ragnar Holm | 715 |
| **Study of the e.m.f. Method** C. T. Tai | 717 |
| **Book Review** | 723 |
| **Letters to the Editor:** | |
| **Ionospheric Absorption** R. N. Bracewell and K. Weekes | 724 |
| **Errata: On Theoretical Signal-to-Noise Ratios in F-M Receivers: A Comparison with Amplitude Modulation** David Middleton | 724 |
| **Erratum: The Thermal Behavior of Evaporated Layers in Vacuum Devices** H. A. Stahl | 724 |
| **Index to Advertiser** | XXIV |



**BOARD OF EDITORS**
Elmer Hutchisson, *Editor*

*Associate Editors*
R. Bowling Barnes
Otto Beeck
W. James Lyons
Philip Morrison
Paul E. Sabine
Raymond J. Seeger
Frederick Seitz
Charles S. Smith
J. R. Townsend

**Publication of the American Institute of Physics**

**OFFICERS**
George R. Harrison, *Chairman*
Wallace Waterfall, *Secretary*
George B. Pegram, *Treasurer*
John T. Tate, *Adviser on Publications*

**GOVERNING BOARD**
Luis W. Alvarez
J. W. Beams
Wallace R. Brode
E. U. Condon
W. F. Fair, Jr.
Harvey Fletcher
George R. Harrison
W. V. Houston
Paul Kirkpatrick
Paul E. Klopsteg
F. W. Loomis
Philip M. Morse
J. R. Oppenheimer
George B. Pegram
Harold K. Schilling
John T. Tate
Wallace Waterfall
M. W. White
A. G. Worthing

**Consultant on Printing**
Melvin Loos

**ADMINISTRATIVE STAFF**
Henry A. Barton, *Director*
Theodore Vorburger, *Business Manager*
Ruth F. Bryans, *Publication Manager*
Betty Hofstadter, *Office Manager*
Eileen B. Neuberger, *Circulation Manager*



The Journal of Applied Physics, published monthly at Prince and Lemon Sts., Lancaster, Pa., is devoted to physics in its role as the science basic to other natural sciences and to the arts and industries. Previous to 1937 this publication was known as Physics. The Journal of Applied Physics publishes editorials, news of physicists, and reviews, as well as technical papers of applied physics. Articles appearing in it are indexed regularly, according to their field, by Physics Abstracts, Chemical Abstracts, Engineering Index, Inc., Industrial Arts Index, etc.

**Manuscripts** should be submitted to Elmer Hutchisson, Editor, Case Institute of Technology, Cleveland 6, Ohio. The authors' institutions are requested to pay a **publication charge** of $4 per page which, if honored, entitles them to 100 free reprints. Instructions will be sent with galley proofs.

**Subscriptions** and orders for back numbers should be addressed to Prince and Lemon Streets, Lancaster, Pa. or to the American Institute of Physics, 57 East 55 Street, New York 22, N. Y.

| **Subscription Price** | U.S. and Canada | Elsewhere |
|---|---|---|
| To members and associate members of the American Institute of Physics | $7.00 | $ 8.00 |
| To all others | 9.00 | 10.00 |

**Back Number Prices**
Complete set: Vol. 1, July, 1931–Vol. 19, 1948—$140.00
Yearly back number rate: $10.00
Single copies: $1.00 each.

**Changes of address** should be sent to the Business Manager, American Institute of Physics; **corrected proofs** should be sent to the Publication Manager, American Institute of Physics.

**Advertising rates** supplied on request. Orders, advertising copy, and cuts should be sent to the American Institute of Physics.

Entered as second class matter January 22, 1937, at the Post Office at Lancaster, Pennsylvania, under the Act of March 3, 1879.
Accepted for mailing at the special rate of postage provided for in the Act of February 28, 1925, authorized May 2, 1932.



Copyright 1949, by the American Institute of Physics.

Journal
of
Applied Physics

Volume 20, Number 7 July, 1949

# Compressibility Effects in Two-Phase Flow

R. F. Tangren, C. H. Dodge, and H. S. Seifert
*Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California*
(Received September 13, 1948)

A gas-water mixture, when expanded through a de Laval nozzle, acts as a compressible fluid. A general expression for the flow process in an idealized mixture is developed. The general thermohydrodynamic relations for flow through a de Laval nozzle are developed. The magnitude of the velocity of an energy pulse or signal through the mixture is determined and is compared with the velocity of flow. Experimental results are compared with those predicted using the equations developed.

## I. INTRODUCTION

THE behavior of a two-phase system, such as a mixture of small gas bubbles in water, has attracted interest for a considerable number of years. As early as 1930 Ackeret[1] derived an expression for the velocity of sound in a gas-water mixture, and the dynamics of such systems has been studied by Heinrich[2] (1942), Charyk[3] (1943), and Zwicky[4] (1944). References to publications of these men, as well as more recent workers in the field, are given in the bibliography, Section IX of this paper. Such a gas-water system is of interest because of the possibility of applying it to the propulsion of underwater devices in which a gas is injected into water, the working fluid.

The procedure used in the analysis to follow is to apply the basic laws of continuity, momentum, energy, and ideal gas equations of state to a mixture in such a way as to derive:

(a) an equation of state for the mixture,
(b) an equation of motion for the mixture.

These equations are then examined for the special cases in which the mixture approaches:

(a) the all gas state,
(b) the all liquid state,
(c) an intermediate state corresponding to a practical experimental mass ratio.

The flow of the "practical" mixture through a de Laval nozzle is calculated and plotted and some comparisons are made with experimental measurements.

The present analysis is greatly simplified by the assumptions tabulated below. The results, therefore, are only indicative of what might follow from a more detailed consideration involving viscosity, turbulence, heat transfer, surface tension, vapor pressure, resonance, and the like. Surprisingly good correlation is found in spite of the simplicity of these assumptions. It is assumed that:

1. The liquid is an incompressible fluid, and effects due to its viscosity, surface tension, and vapor pressure are unimportant.

2. The gas is an ideal gas, with negligible viscosity, constant specific heats, and is insoluble in the liquid.

3. The mixture is "homogeneous" in the sense that the bubbles of gas are so small and uniformly distributed that an arbitrarily small sample contains the same mass ratio of gas to liquid as the whole mixture.

4. The gas and liquid are always at the same

[1] J. Ackeret, "Experimental and theoretical investigations on cavitation in water" (Translation from *Technische Mechanik und Thermodynamik* (Forschung), Vol. 1, No. 1, Berlin, 1930).
[2] G. Heinrich, "Über Stromungen von Schaumen," Zeits. f. angew. Math. Mech. 22, 117–118 (1942).
[3] J. Charyk, Jet Propulsion Laboratory Progress Report 2–2 (California Institute of Technology, November, 1943).
[4] F. Zwicky, Aerojet Engineering Corporation Report R-10 (January, 1944).

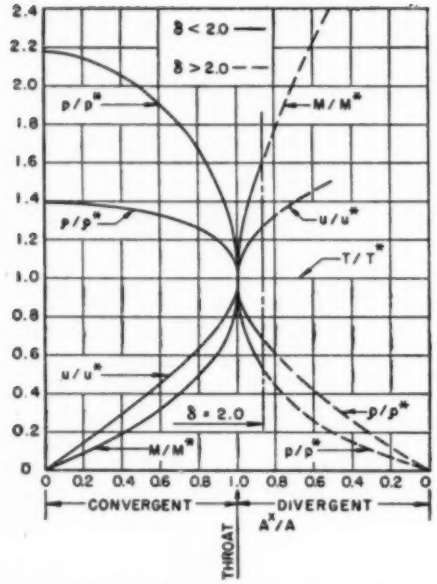


FIG. 1. Dimensionless graph of supersonic nozzle flow and thermodynamic parameters, referred to values at throat as functions of the relative area and for the particular value of $\delta_0 = 0.50$.

temperature, and the flow is insulated (i.e., adiabatic).

5. The flow is one-dimensional, laminar, and expands or compresses slowly enough that inertial transients may be ignored.

6. A *slowly varying* pressure change or signal is transmitted through the mixture with a definite critical or "sonic" velocity, as distinguished from the anomalous dispersion and attenuation of high frequency pressure waves.

The nomenclature is defined here in alphabetical order for reference. Symbols will be redefined in the text at the places where they first appear.

### Roman Letters

$a$—sonic or critical velocity, ft./sec.
$A$—area, ft.$^2$.
$C_f$—specific heat of liquid, ft.$^2$/sec.$^2$°$R$.
$C_p$—specific heat of gas at constant pressure, ft.$^2$/sec.$^2$°$R$ or ft. lb./slug°$R$.
$C_v$—specific heat of gas at constant volume, ft.$^2$/sec.$^2$°$R$.
$M_f$—mass of liquid, slugs per ft.$^3$ of mixture.
$M_g$—Mass of gas, slugs per ft.$^3$ of mixture.
$p$—pressure, lb./ft.$^2$.
$Q$—heat added per slug, ft.$^2$/sec.$^2$.
$R$—gas constant, ft.$^2$/sec.$^2$°$R$.
$T$—temperature, °$R$.
$u$—velocity, ft./sec.
$W$—work done per slug, ft.$^2$/sec.$^2$.
$()_0$—value under initial conditions.
$()^*$—conditions at the throat.

### Greek Letters

$\gamma$—ratio of specific heat of gas at constant pressure to the specific heat of gas at constant volume.
$\delta$—ratio of volume of gas to volume of liquid per unit volume of mixture.
$\mu$—ratio of mass of gas to mass of liquid per unit volume of mixture.
$\rho$—mass density of mixture, slugs/ft.$^3$.
$\rho_f$—fluid mass density, slugs/ft.$^3$.
$\rho_g$—gas mass density, slugs/ft.$^3$.

## II. EQUATION OF STATE FOR A TWO-PHASE FLUID

A unit volume of *mixture* at any instant is occupied by a mass $M_g$ of gas and $M_f$ of liquid, which are in the *mass ratio*

$$\mu = M_g/M_f.$$

Since the volume of gas plus the volume of liquid must equal the total volume, this fact may be used to define an average *mixture density* as follows:

$$M_g/\rho_g + M_f/\rho_f = (M_g + M_f)/\rho. \quad (1)$$

As the mixture expands or contracts, $\mu$ remains fixed, but the *volume ratio*

$$\delta = \mu\rho_f/\rho_g$$

varies inversely with the gas density. By combining the definitions of $\mu$ and $\delta$ with Eq. (1) above, one arrives at two useful relations among the densities

$$\rho/\rho_f = (1+\mu)/(1+\delta) \quad (2)$$

and

$$1/\rho_g = (1/\mu)\left[\frac{1+\mu}{\rho} - \frac{1}{\rho_f}\right]. \quad (3)$$

The equation for conservation of energy, when written in terms of energy per unit *volume*, with the assumption that gas and liquid temperatures are

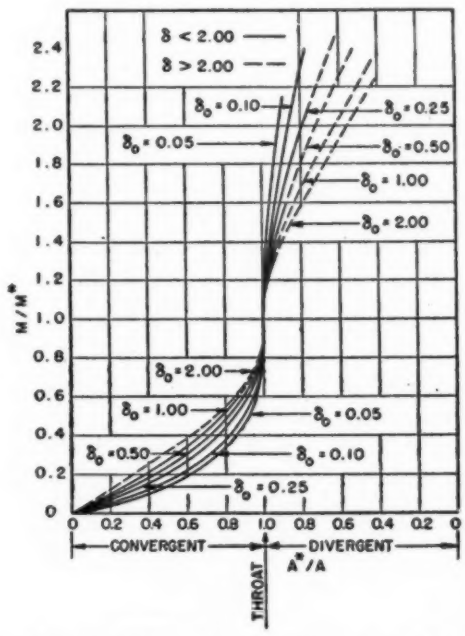


FIG. 2. Relation between Mach number ratio ($M/M^*$) and area ratio ($A^*/A$) for various values of $\delta_0$.

always equal, takes the form

$$(M_g+M_f)dQ=(M_gC_v+M_fC_f)dT+(M_g+M_f)dW \quad (4)$$

or

$$dQ=[(\mu C_v+C_f)/(1+\mu)]dT+dW. \quad (5)$$

The work done in expanding a unit mass of the fluid against the external pressure $p$ is $dW=pd(1/\rho)$ and if this is substituted in Eq. (5), together with the fact that in an insulated flow $dQ=0$, we have

$$(p/\rho^2)d\rho=[(\mu C_v+C_f)/(1+\mu)]dT. \quad (6)$$

The equation of state for the gas component alone is

$$p/\rho_g=RT. \quad (7)$$

One makes use of the equation of state (7) and the density relation (3) to eliminate the unwanted variables $\rho_g$ and $T$ and to arrive at the differential equation of state involving only pressure $p$ and mixture density $\rho$ as variables. Inserting (3) in (7), differentiating, and combining the result with (6) yields the adiabatic state equation

$$dp/p=[(\mu C_v+C_f+\mu R)/(\mu C_v+C_f)]\times\left[d\rho/\rho\left(1-\frac{1}{1+\mu}\frac{\rho}{\rho_f}\right)\right], \quad (8)$$

which may be integrated (noting that $C_p=C_v+R$) to

$$p[(1/\rho)-1/\rho_f(1+\mu)]^{(\mu C_p+C_f)/(\mu C_v+C_f)}=\text{constant}. \quad (9)$$

If the new and significant parameter is defined

$$\Gamma=(\mu C_p+C_f)/(\mu C_v+C_f)$$

and Eq. (9) simplified with $\Gamma$ and the density

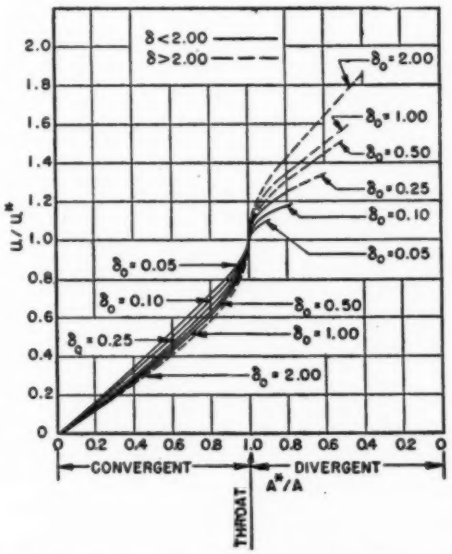


FIG. 3. Relation between velocity ratio ($u/u^*$) and area ratio ($A^*/A$) for various values of $\delta_0$.

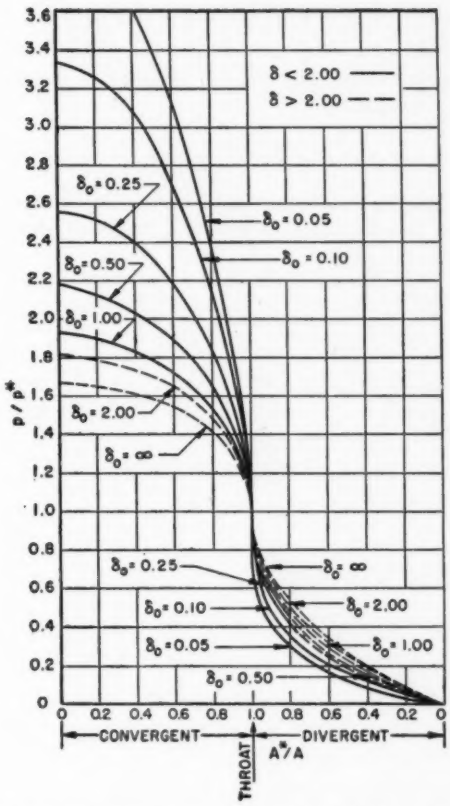


FIG. 4. Relation between pressure ratio ($p/p^*$) and area ratio ($A^*/A$) for various values of $\delta_0$.

equation (3), then

$$p[\mu/(1+\mu)\rho_g]^\Gamma=\text{constant}. \quad (10)$$

To express the adiabatic state equation in terms of temperature, insert (7) in Eq. (10), from which

$$p[\{\mu/(1+\mu)\}\{RT/p\}]^\Gamma=\text{constant} \quad (11)$$

or

$$p^{(\Gamma-1)/\Gamma}/T=\text{constant}. \quad (12)$$

It may be seen from Eq. (12) that $\Gamma$ has a significance for adiabatic mixed flow similar to that of $\gamma$ in adiabatic gas flow.

From the definition of $\Gamma$ as $\mu\to\infty$

$$\Gamma\to C_p/C_v=\gamma.$$

Thus the limiting condition in which $\mu\to\infty$, (pure gas) gives the familiar adiabatic equation for a gas

$$p^{(\gamma-1)/\gamma}/T=\text{constant}.$$

On the other hand, as $\mu\to 0$ (pure liquid) $\Gamma\to 1.0$ and the process is isothermal with $T=\text{constant}$.

### III. EQUATION OF MOTION FOR A TWO-PHASE FLUID

By applying to Newton's equation of motion the equation of state just derived (Eq. (9)), there results an expression stating bulk fluid velocity $u$ in

terms of pressure and known fixed quantities. The assumption of no viscosity and one-dimensional flow permits the use of the following simple form of Euler's equation:

$$udu = -dp/\rho. \tag{13}$$

Eliminating $1/\rho$ by means of Eq. (9) and integrating between the limits $u$, $p$ and $u_0$, $p_0$ gives the general result

$$u^2 = u_0^2 + [2\Gamma/(\Gamma-1)][\mu/(1+\mu)][1/\rho_{g0}] \times [p_0 - p(p_0/p)^{1/\Gamma}] + 2(p_0-p)/\rho_f(1+\mu). \tag{14}$$

Equation (14) may be expressed in terms of volume ratio $\delta_0$ (where the subscript 0 refers to the initial point of the flow) by using the relation, derivable from Eq. (10), that

$$(p/p_0)^{1/\Gamma} = \delta_0/\delta. \tag{15}$$

Hence, we have finally

$$u^2 = u_0^2 + 2p_0/\rho_f(1+\mu)\{[\Gamma/(\Gamma-1)]\delta_0 \times [1-(p/p_0)^{(\Gamma-1)/\Gamma}] + (1-p/p_0)\}. \tag{16}$$

It is of interest to examine the general relation (16) for the cases of pure gas and pure liquid, respectively. By definition, $\delta_0 = \mu\rho_f/\rho_{g0}$. If we use this and allow $\mu \to \infty$, then (16) becomes

$$u^2 = u_0^2 + 2(p_0/\rho_{g0})[\gamma/(\gamma-1)] \times [1-(p/p_0)^{(\gamma-1)/\gamma}], \tag{17}$$

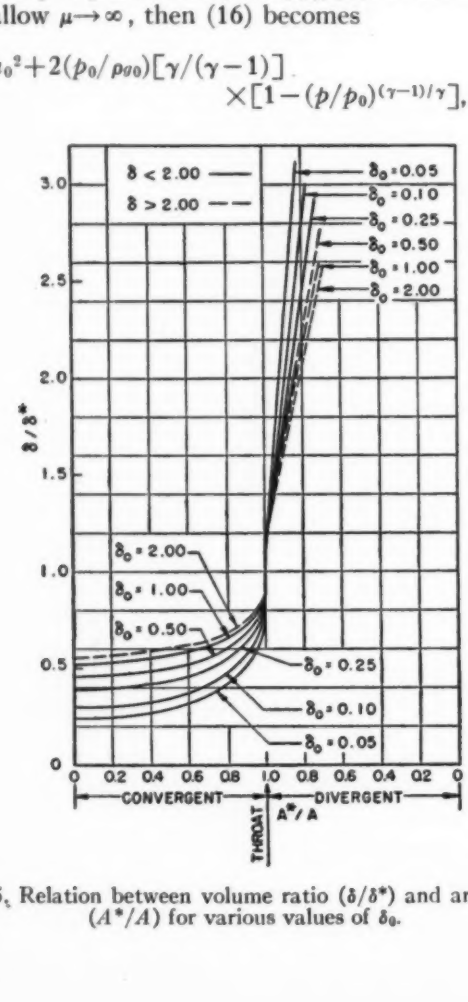


FIG. 5. Relation between volume ratio ($\delta/\delta^*$) and area ratio ($A^*/A$) for various values of $\delta_0$.

which is the familiar equation for the adiabatic expansion of a pure gas.

If we let $\mu \to 0$, Eq. (16) reduces to the well-known Bernoulli equation for an incompressible fluid

$$u^2 = u_0^2 + 2(p_0/\rho_f)[1-p/p_0]. \tag{18}$$

## IV. FLOW OF A TYPICAL MIXTURE

In order to continue the analysis of mixture flow in a quantitative manner, a particular practical case will be considered. This makes it possible to take advantage of certain simplifying approximations. The example now to be discussed is the water-helium system, accelerating from rest.

In order that the gas bubbles be able to do work during expansion on the liquid matrix in which they are imbedded, the ratio of gas to liquid volume must be limited to some maximum value, or the liquid will become dispersed in such a way that it no longer encloses the gas. The volume ratio $\delta$ has been arbitrarily limited to 2.0. It should be noted that this is the *maximum* value of $\delta$ and that $\delta_0$ is usually smaller than this.

Using water as the liquid and helium as the gas, with $\delta = 2.0$, typical values of $\mu$ and $\Gamma$ for several pressures are calculated below:

| $p$ | $\mu$ | $\Gamma$ |
|---|---|---|
| 1 atmos. | $4.4\times10^{-4}$ | 1.00022 |
| 2 atmos. | $8.9\times10^{-4}$ | 1.00044 |
| 3 atmos. | $17.8\times10^{-4}$ | 1.00088 |

At the end of Section II it was shown that when $\mu$ is very small, $\Gamma$ approaches unity and the mixture behaves as if the gas process were isothermal. It is seen from the above table that this is the case for a practical gas-water mixture. Physically this means that the mass and specific heat of the liquid are so large relative to the gas that the gas may remain in thermal equilibrium with the liquid during expansion without appreciably lowering the liquid temperature. This is certainly a close approximation so long as $\delta < 2.0$.

In view of the fact that $\Gamma \to 1.0$ and $\mu \to 0$, Eq. (16) may be somewhat simplified. The expression $(p/p_0)^{(\Gamma-1)/\Gamma}$ can be expanded in an exponential series, of which only the first two terms need be used.

$$(p/p_0)^{(\Gamma-1)/\Gamma} \doteq 1 + [(\Gamma-1)/\Gamma]\ln p/p_0.$$

Neglecting $\mu$ relative to unity in (16) gives

$$u^2 \doteq u_0^2 - 2(p_0/\rho_f)\delta_0(\ln p/p_0) + 2p_0/\rho_f[1-p/p_0]$$

to a close approximation.

Assuming expansion from rest, $u_0 = 0$, we may write a dimensionless form of the velocity equation,

$$\rho_f u^2/2p_0 \doteq -\delta_0(\ln p/p_0) + 1 - p/p_0. \tag{19}$$

Because of the approximations involved, Eq. (19) indicates that as the free stream pressure $p$ approaches zero, the velocity $u$ approaches infinity. This would be true if the liquid acted as an infinite heat source and maintained the gas expansion strictly isothermal. In an actual physical case, $\delta$ would exceed the limiting value of 2.0 long before the heat of the liquid was perceptibly depleted.

It is interesting to note in Eq. (19) that the only property of the gas which enters into the calculation of velocity is its volume ratio $\delta_0$. In other words, the molecular weight, specific heat, etc., do not affect $u$ in this approximate case, and a given volume of one species of ideal gas is as effective as any other.

## V. VELOCITY OF A PRESSURE WAVE THROUGH A MIXTURE

A basic assumption, made at the beginning of this analysis, was that a slowly varying pressure pulse is transmitted through a mixture with a perfectly definite "sonic" velocity. If this be so, then it is possible to apply the results of de Laval nozzle theory and show that the bulk velocity of the fluid equals this "sonic" velocity at the throat of a nozzle through which a mixture is flowing, that is, the Mach number is unity at the throat.

It is realized that ordinary sound waves or *rapidly* fluctuating pressures are transmitted through bubbles with a complicated array of dispersion, attenuation, scattering, and other phenomena. It is assumed here that the pressure variations existing in nozzle flow, unlike high frequency sound waves, are transmitted as if the medium were essentially homogeneous.

It can be shown[5] that the basic equation for the velocity $a$ of a wave or signal through an isotropic homogeneous elastic medium is

$$a^2 = dp/d\rho.$$

The term $dp/d\rho$ may be evaluated from the previously derived equation of state, Eq. (8). By operating on Eq. (8) with Eqs. (2) and (3) and the definitions of $\delta$ and $\delta_0$, one may obtain the expression for the speed of a signal through a mixture as

$$a^2 = (\Gamma p/\rho_0)[\mu/(1+\mu)][1+1/\delta]^2. \qquad (20)$$

Equation (20) may be verified for extreme values of $\mu$. As $\mu \to \infty$

$$a^2 \to \gamma p/\rho_0 = \gamma RT,$$

which is the equation for the adiabatic speed of sound in the gas. As $\mu \to 0$, $a^2 \to \infty$, or the speed of sound in an incompressible fluid.

One may write an approximate form of Eq. (20)

[5] A. B. Wood, *A Textbook of Sound* (G. Bell and Sons, London, 1930), p. 51.

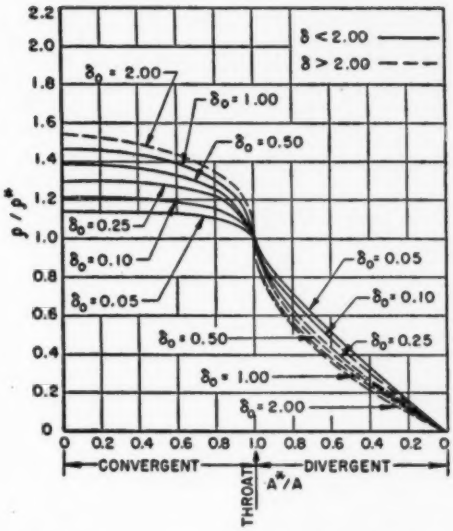


FIG. 6. Relation between density ratio $(\rho/\rho^*)$ and area ratio $(A^*/A)$ for various values of $\delta_0$.

for the practical case where $\mu \to 0$ and $\Gamma \to 1.0$ as follows:

$$a^2 \doteq (p_0\delta_0/\rho_f)[1+1/\delta]^2. \qquad (21)$$

With the help of Eq. (15), in which the approximation $\Gamma \to 1.0$ has also been used, Eq. (21) may be put into the same form as Eq. (19):

$$\rho_f a^2/2p_0 \doteq (\delta_0/2)[1+p/p_0\delta_0]^2. \qquad (22)$$

## VI. FLOW THROUGH A DE LAVAL NOZZLE

If the mixture flow now be considered to be confined to a duct of varying cross section, the equation of continuity requires that

$$dA/A + d\rho/\rho + du/u = 0.$$

If the duct has a minimum cross section or throat, as in the de Laval nozzle, at that point

$$dA/A = 0.$$

Combining the continuity equation (with the throat condition) with Euler's Eq. (13), one has the basic result that the velocity of flow *at the throat* must be

$$u^2 = dp/d\rho.$$

This, however, is just the magnitude of the "sonic" velocity as stated in a preceding section. It is thus concluded that the *flow velocity at the throat equals the local signal velocity*. This "critical" velocity is denoted by a superscript *, and at the throat

$$u = a = u^* = a^*.$$

In order to evaluate as a function of $\delta_0$ the critical pressure ratio $p^*/p_0$, at which sonic flow velocity is achieved, the approximate expression (22) for $a^2$ is equated to the expression (19) for $u^2$. Since

TABLE I. Nozzle flow parameters. $\delta_0$: initial ratio of the mass of gas to the mass of the fluid per unit volume of the mixture. $p/p_0$: ratio of the pressure at any point to the initial pressure. $p^*/p_0$: ratio of the pressure at the throat to the initial pressure. $u/(2p_0/\rho_f)^{\frac{1}{2}}$: non-dimensional velocity at any point. $u^*/(2p_0/\rho_f)^{\frac{1}{2}}$: non-dimensional velocity at the throat. $a/(2p_0/\rho_f)^{\frac{1}{2}}$: non-dimensional speed of sound at any point. $M$: Mach number, or ratio of velocity to the speed of sound. $\delta/\delta^*$: ratio of the mass of the gas to the mass of the fluid ratio at any point to that at the throat. $\rho/\rho^*$: ratio of the density at any point to that at the throat. $A^*/A$: ratio of the nozzle throat cross-sectional area to the duct cross-sectional area at any point.

| $\delta_0$ | $p/p_0$ | $u/(2p_0/\rho_f)^{\frac{1}{2}}$ | $a/(2p_0/\rho_f)^{\frac{1}{2}}$ | $M$ | $u^*/(2p_0/\rho_f)^{\frac{1}{2}}$ | $u/u^*$ | $p^*/p_0$ | $p/p^*$ | $\delta/\delta^*$ | $\rho/\rho^*$ | $A^*/A$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.0 | 0.10 | 2.35 | 1.05 | 2.24 | 1.28 | 1.84 | 0.55 | 0.18 | 5.50 | 0.22 | 0.41 |
| 2.0 | 0.20 | 2.00 | 1.10 | 1.83 | 1.28 | 1.57 | 0.55 | 0.36 | 2.75 | 0.42 | 0.67 |
| 2.0 | 0.30 | 1.77 | 1.15 | 1.54 | 1.28 | 1.38 | 0.55 | 0.54 | 1.83 | 0.61 | 0.84 |
| 2.0 | 0.40 | 1.56 | 1.20 | 1.30 | 1.28 | 1.22 | 0.55 | 0.73 | 1.37 | 0.77 | 0.95 |
| 2.0 | 0.50 | 1.37 | 1.25 | 1.09 | 1.28 | 1.07 | 0.55 | 0.91 | 1.10 | 0.93 | 1.00 |
| 2.0 | 0.60 | 1.19 | 1.30 | 0.92 | 1.28 | 0.93 | 0.55 | 1.09 | 0.92 | 1.07 | 1.00 |
| 2.0 | 0.70 | 1.01 | 1.35 | 0.75 | 1.28 | 0.79 | 0.55 | 1.27 | 0.79 | 1.20 | 0.95 |
| 2.0 | 0.80 | 0.81 | 1.40 | 0.57 | 1.28 | 0.63 | 0.55 | 1.45 | 0.69 | 1.33 | 0.84 |
| 2.0 | 0.90 | 0.56 | 1.45 | 0.38 | 1.28 | 0.44 | 0.55 | 1.63 | 0.63 | 1.44 | 0.62 |
| 2.0 | 1.00 | — | — | — | 1.28 | — | 0.55 | 1.81 | 0.55 | 1.55 | — |
| 1.0 | 0.10 | 1.80 | 0.78 | 2.31 | 1.07 | 1.68 | 0.52 | 0.19 | 5.17 | 0.27 | 0.45 |
| 1.0 | 0.20 | 1.55 | 0.85 | 1.82 | 1.07 | 1.45 | 0.52 | 0.39 | 2.60 | 0.49 | 0.71 |
| 1.0 | 0.30 | 1.38 | 0.92 | 1.50 | 1.07 | 1.29 | 0.52 | 0.58 | 1.73 | 0.68 | 0.88 |
| 1.0 | 0.40 | 1.23 | 0.99 | 1.24 | 1.07 | 1.15 | 0.52 | 0.77 | 1.30 | 0.84 | 0.97 |
| 1.0 | 0.50 | 1.09 | 1.06 | 1.03 | 1.07 | 1.02 | 0.52 | 0.97 | 1.03 | 0.98 | 1.00 |
| 1.0 | 0.60 | 0.96 | 1.13 | 0.85 | 1.07 | 0.89 | 0.52 | 1.16 | 0.87 | 1.10 | 0.98 |
| 1.0 | 0.70 | 0.81 | 1.20 | 0.68 | 1.07 | 0.76 | 0.52 | 1.35 | 0.74 | 1.21 | 0.92 |
| 1.0 | 0.80 | 0.65 | 1.27 | 0.51 | 1.07 | 0.61 | 0.52 | 1.55 | 0.65 | 1.30 | 0.79 |
| 1.0 | 0.90 | 0.45 | 1.34 | 0.34 | 1.07 | 0.42 | 0.52 | 1.74 | 0.58 | 1.39 | 0.59 |
| 1.0 | 1.00 | — | — | — | 1.07 | — | 0.52 | 1.94 | 0.52 | 1.47 | — |
| 0.50 | 0.10 | 1.43 | 0.60 | 2.39 | 0.96 | 1.50 | 0.46 | 0.22 | 4.59 | 0.35 | 0.52 |
| 0.50 | 0.20 | 1.27 | 0.70 | 1.82 | 0.96 | 1.33 | 0.46 | 0.44 | 2.30 | 0.60 | 0.79 |
| 0.50 | 0.30 | 1.14 | 0.80 | 1.43 | 0.96 | 1.19 | 0.46 | 0.65 | 1.53 | 0.78 | 0.93 |
| 0.50 | 0.40 | 1.03 | 0.90 | 1.15 | 0.96 | 1.08 | 0.46 | 0.87 | 1.15 | 0.93 | 1.00 |
| 0.50 | 0.50 | 0.92 | 1.00 | 0.92 | 0.96 | 0.96 | 0.46 | 1.09 | 0.92 | 1.04 | 1.00 |
| 0.50 | 0.60 | 0.81 | 1.10 | 0.74 | 0.96 | 0.85 | 0.46 | 1.31 | 0.77 | 1.14 | 0.96 |
| 0.50 | 0.70 | 0.69 | 1.20 | 0.58 | 0.96 | 0.72 | 0.46 | 1.52 | 0.66 | 1.22 | 0.88 |
| 0.50 | 0.80 | 0.56 | 1.30 | 0.43 | 0.96 | 0.58 | 0.46 | 1.74 | 0.57 | 1.28 | 0.75 |
| 0.50 | 0.90 | 0.39 | 1.40 | 0.28 | 0.96 | 0.41 | 0.46 | 1.96 | 0.51 | 1.34 | 0.55 |
| 0.50 | 1.00 | — | — | — | 0.96 | — | 0.46 | 2.18 | 0.46 | 1.39 | — |
| 0.25 | 0.10 | 1.22 | 0.50 | 2.46 | 0.92 | 1.33 | 0.39 | 0.26 | 3.92 | 0.46 | 0.62 |
| 0.25 | 0.20 | 1.10 | 0.64 | 1.73 | 0.92 | 1.20 | 0.39 | 0.51 | 1.96 | 0.72 | 0.88 |
| 0.25 | 0.30 | 1.00 | 0.78 | 1.28 | 0.92 | 1.09 | 0.39 | 0.77 | 1.31 | 0.89 | 0.98 |
| 0.25 | 0.40 | 0.91 | 0.92 | 0.99 | 0.92 | 1.00 | 0.39 | 1.03 | 0.98 | 1.00 | 1.00 |
| 0.25 | 0.50 | 0.82 | 1.06 | 0.77 | 0.92 | 0.90 | 0.39 | 1.28 | 0.78 | 1.08 | 0.98 |
| 0.25 | 0.60 | 0.73 | 1.21 | 0.60 | 0.92 | 0.79 | 0.39 | 1.54 | 0.65 | 1.15 | 0.92 |
| 0.25 | 0.70 | 0.62 | 1.35 | 0.46 | 0.92 | 0.68 | 0.39 | 1.80 | 0.56 | 1.20 | 0.83 |
| 0.25 | 0.80 | 0.51 | 1.49 | 0.34 | 0.92 | 0.55 | 0.39 | 2.05 | 0.49 | 1.24 | 0.69 |
| 0.25 | 0.90 | 0.36 | 1.63 | 0.22 | 0.92 | 0.39 | 0.39 | 2.31 | 0.44 | 1.27 | 0.50 |
| 0.25 | 1.00 | — | — | — | 0.92 | — | 0.39 | 2.57 | 0.39 | 1.30 | — |
| 0.10 | 0.10 | 1.06 | 0.45 | 2.38 | 0.90 | 1.18 | 0.30 | 0.33 | 3.00 | 0.67 | 0.78 |
| 0.10 | 0.20 | 0.98 | 0.67 | 1.46 | 0.90 | 1.09 | 0.30 | 0.67 | 1.52 | 0.89 | 0.96 |
| 0.10 | 0.30 | 0.91 | 0.90 | 1.02 | 0.90 | 1.00 | 0.30 | 1.00 | 1.00 | 1.00 | 1.00 |
| 0.10 | 0.40 | 0.83 | 1.12 | 0.74 | 0.90 | 0.92 | 0.30 | 1.33 | 0.76 | 1.07 | 0.98 |
| 0.10 | 0.50 | 0.76 | 1.34 | 0.56 | 0.90 | 0.84 | 0.30 | 1.67 | 0.60 | 1.11 | 0.92 |
| 0.10 | 0.60 | 0.67 | 1.57 | 0.43 | 0.90 | 0.75 | 0.30 | 2.00 | 0.51 | 1.14 | 0.85 |
| 0.10 | 0.70 | 0.58 | 1.79 | 0.32 | 0.90 | 0.64 | 0.30 | 2.33 | 0.43 | 1.17 | 0.75 |
| 0.10 | 0.80 | 0.47 | 2.02 | 0.23 | 0.90 | 0.52 | 0.30 | 2.67 | 0.38 | 1.19 | 0.62 |
| 0.10 | 0.90 | 0.33 | 2.24 | 0.15 | 0.90 | 0.37 | 0.30 | 3.00 | 0.34 | 1.20 | 0.44 |
| 0.10 | 1.00 | — | — | — | 0.90 | — | 0.30 | 3.33 | 0.30 | 1.21 | — |
| 0.05 | 0.10 | 1.01 | 0.47 | 2.13 | 0.92 | 1.10 | 0.25 | 0.40 | 2.38 | 0.80 | 0.89 |
| 0.05 | 0.20 | 0.94 | 0.79 | 1.19 | 0.92 | 1.03 | 0.25 | 0.80 | 1.19 | 0.96 | 0.99 |
| 0.05 | 0.30 | 0.87 | 1.11 | 0.79 | 0.92 | 0.96 | 0.25 | 1.20 | 0.80 | 1.03 | 0.99 |
| 0.05 | 0.40 | 0.81 | 1.42 | 0.57 | 0.92 | 0.88 | 0.25 | 1.60 | 0.60 | 1.07 | 0.94 |
| 0.05 | 0.50 | 0.73 | 1.74 | 0.42 | 0.92 | 0.80 | 0.25 | 2.00 | 0.48 | 1.09 | 0.88 |
| 0.05 | 0.60 | 0.65 | 2.06 | 0.32 | 0.92 | 0.72 | 0.25 | 2.40 | 0.40 | 1.11 | 0.79 |
| 0.05 | 0.70 | 0.57 | 2.36 | 0.24 | 0.92 | 0.62 | 0.25 | 2.80 | 0.34 | 1.12 | 0.70 |
| 0.05 | 0.80 | 0.46 | 2.69 | 0.17 | 0.92 | 0.50 | 0.25 | 3.20 | 0.30 | 1.13 | 0.57 |
| 0.05 | 0.90 | 0.32 | 3.01 | 0.11 | 0.92 | 0.35 | 0.25 | 3.60 | 0.27 | 1.14 | 0.40 |
| 0.05 | 1.00 | — | — | — | 0.92 | — | 0.25 | 4.00 | 0.24 | 1.14 | — |

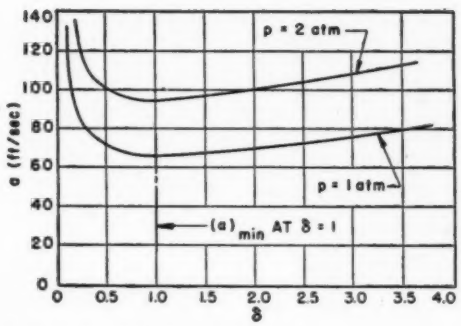


FIG. 7. Relation between velocity of a pressure wave ($a$) and $\delta$ with constant pressure ($p$).

$$u^{*2}=a^{*2},$$

$$(\delta_0/2)(1+p^*/p_0\delta_0)^2 \doteq (1-p^*/p_0)-\delta_0 \ln p^*/p_0. \quad (23)$$

Equation (23) may be solved graphically to give the pressure ratio corresponding to "sonic" velocity as a function of the initial volume ratio $\delta_0$. This equation is limited in accuracy by the approximation $\mu \to 0$ and $\Gamma \to 1.0$.

## VII. GRAPHICAL PRESENTATION OF NOZZLE FLOW PARAMETERS

The equations derived in the preceding sections are sufficient to supply data on the variation of flow velocity, sonic velocity, area, density, gas volume, and pressure throughout a de Laval nozzle. Most of the parameters have been expressed in terms of pressure ratio $p/p_0$ as the independent variable. It will be more instructive and convenient, however, to plot them as a function of the nozzle area ratio $A/A^*$ as the independent variable. This is done by calculating $A/A^*$ as a function of $p/p_0$ and plotting all quantities against area rather than pressure.

In dimensionless form the quantities of interest are (note that * refers to throat conditions):

| | |
|---|---|
| Local Mach number ratio.. | $M/M^*=u/a$. |
| Velocity relative to throat. | $u/u^*$. |
| Pressure ratio..... | $p/p^*$. |
| Volume ratio..... | $\delta/\delta^*$. |
| Density ratio..... | $\rho/\rho^*$. |
| Area ratio..... | $A/A^*$. |

These quantities are plotted together in Fig. 1 for a particular value of $\delta_0$ and separately for a range of values of $\delta_0$ in Figs. 2–6. The sequence of calculations necessary to arrive at these curves are now described in some detail. The numerical values have been tabulated in Table I.

### A. Local Mach Number $u/a$

The ratio of $u/a$ may be determined by computing $u/(2p_0/\rho_f)^{\frac{1}{2}}$ and $a/(2p_0/\rho_f)^{\frac{1}{2}}$ from Eqs. (19) and (22).

This ratio is the local Mach number $M$, and since the throat Mach number $M^*=1$, $M=M/M^*$ also. It is computed for a series of values of $p/p_0$ and $\delta_0$. See Fig. 2.

### B. Velocity Relative to Throat Velocity $u/u^*$

Since $u=u^*=a$ at the throat, expressions (19) and (22) for $u$ and $a$ may be equated to determine graphically or numerically the pressure ratio $p^*/p_0$ which exists at the throat; $p^*/p_0$ is, of course, a function of $\delta_0$. We use this value of $p^*/p_0$ in (19) to find $u^*/(2p_0/\rho_f)^{\frac{1}{2}}$, which is then divided into the previously determined $u/(2p_0/\rho_f)^{\frac{1}{2}}$. $u/u^*$ is shown in Fig. 3.

### C. Pressure, Volume, and Density Ratios

Pressure relative to the throat, $p/p^*$, may be simply found once the critical pressure ratio $p^*/p_0$ is known, since $p/p^*=(p/p_0)(p_0/p^*)$. The volume ratio relative to that at the throat, $\delta/\delta^*$, is approximately the reciprocal of $p/p^*$, as may be seen from Eq. (15). The density relative to that at the throat may be found by using Eq. (2) with $\mu$ neglected relative to unity to obtain

$$\rho/\rho^* \doteq (1+\delta^*)/(1+\delta),$$
$$\rho/\rho^* \doteq [1+\delta_0/(p^*/p_0)]/[1+\delta_0/(p/p_0)].$$

### D. Area Ratio $A/A^*$

The area relative to that of the throat may be calculated from the density and velocity ratios using the equation of continuity

$$A\rho u = A^*\rho^* u^*,$$

written in the form

$$A/A^* = (\rho^*/\rho)(u^*/u).$$

The values of $A/A^*$ as a function of $p/p_0$ are correlated with the preceding variables $M$, $u$, $p$, $\delta$, and $\rho$, which are then plotted against *area ratio* in Figs. 2–6.

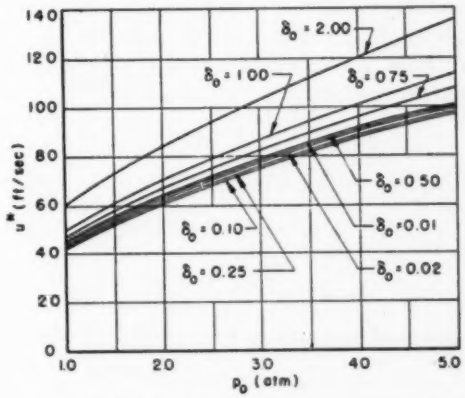


FIG. 8. Relation between $u^*$ and pressure ($p_0$) for various values of $\delta_0$.

FIG. 9. Convergent nozzle with flow of water alone.

FIG. 10. Convergent nozzle with flow of air-water mixture showing expansion after leaving nozzle.

### E. Variation of "Sonic" Speed $a$ with Volume Ratio $\delta$ and Local Pressure $p$

Using Eq. (15), the relation (21) for $a^2$ may be written in the approximate form, valid for $\delta$ positive but less than some small integer:

$$a^2 \doteq (p\delta/\rho_f)[1+1/\delta]^2.$$

It is interesting to note that if the local pressure $p$ is held constant, $a^2$ has a minimum at $\delta=1$, which can be shown by differentiating the above equation with respect to $\delta$. This minimum is surprisingly low. In Fig. 7, $a$ is plotted against $\delta$ for $p=1$ atmos. and $p=2$ atmos., and has a minimum value of 66 ft./sec. at $\delta=1$ for $p=1$ atmos.

It is suggested[6] that the reason for the much lower sonic speed of the mixture when compared with that in either the gas or the liquid is that the gas acts as a weak spring when coupled with a large mass, the liquid. This point can be illustrated by considering the speed of a signal through a pipe containing a liquid-gas mixture. This mixture must be considered homogeneous for this simplified analysis which cannot consider the effects of discontinuities in the medium. The speed of sound through such a tube may be written as $a^2=E/\rho$. Where $E$ is the coefficient of elasticity, and, for the mixture in the tube, $E$ becomes

$$E=-dp/(dV/V).$$

Since for an isothermal gas $dV_g/V_g=-dp/p$, then

$$dV/V=-[\delta/(1+\delta)](dp/p)$$

and

$$E=[(1+\delta)/\delta]p.$$

If the mass of the gas is neglected, $\rho$ becomes

$$\rho=\rho_f/(1+\delta)$$

and the sonic speed squared is

$$a^2=(p/\rho_f)(1+\delta)^2/\delta.$$

Using Eq. (15), considering the process as isothermal ($\Gamma=1$), the sonic speed squared may be written as

$$a^2=(p_0\delta_0/\rho_f)(1+1/\delta)^2.$$

This equation is identical to Eq. (21) derived earlier. Thus it may be seen that this model, combining the elasticity of the gas and the density of the liquid, may be used to obtain a physical picture of the compressibility effects in two-phase flow.

The specific value of the sonic velocity at the throat, $a^*$, is of interest because the actual flow velocity $u^*$ equals $a^*$ at the throat. In Fig. 8, $u^*\equiv a^*$ has been plotted for a range of values of the quantities $p_0$ and $\delta_0$, which are the parameters which are most easily measured and controlled. It should be noted that for any given $p_0$, $u^*$ becomes a minimum at $\delta_0=0.10$, which corresponds to $\delta^*=1.0$.

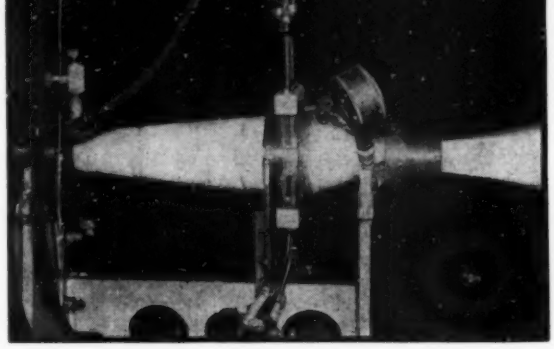

FIG. 11. Mixture flow with convergent-divergent nozzle.

[6] Anderson, Rush, Jr., and McClellan, M.S. Thesis in Aerodynamics, California Institute of Technology, June 1, 1947.

## VIII. COMPARISON OF THEORY AND EXPERIMENT

A limited amount of experimental work was carried out at the Jet Propulsion Laboratory, C. I. T., on the flow of gas-water mixtures through ducts. The object of this work was to study the thrust produced by such a mixture as it expanded through a nozzle.

The gas was injected in a "mixing" section where the duct cross section was large enough that the liquid velocity was negligible. Figure 9 shows the test equipment with water alone flowing through it and Fig. 10 shows the same equipment with gas being injected. Since this duct had only a contracting section, considerable expansion was evident after the mixture left the duct at the point of narrowest cross section. In Fig. 11 a short expanding section had been added to the duct.

So long as incompressible liquid only is flowing through the duct into the atmosphere, the boundary conditions require that the static pressure at the exit of the duct be atmospheric. If gas bubbles are added, the mixture becomes compressible and above the critical flow speed the medium becomes incapable of transmitting pressure changes upstream against the flow. At this point all pressures inside the duct are proportional to "reservoir" pressure $p_0$ and are no longer influenced by the atmospheric, or discharge, pressure. As a consequence, pressures higher than atmospheric may be obtained at the exit section of the nozzle, and expansion to atmospheric pressure may occur after leaving the nozzle, as is implied in Fig. 10. This figure is therefore taken as qualitative evidence that compressible flow at the critical speed is occurring.

A series of runs was made in which increasing amounts of gas were added to a constant volume flow of water, thus increasing both $\delta_0$ and $p_0$. Although it was desired to measure $p_0/p^*$, the chamber to throat pressure ratio, it was actually necessary to measure $p_0/p_1$, where $p_1$ was the pressure at a point slightly upstream from the minimum cross section such that $A^*/A_1$ was 0.993.

TABLE II. Comparison of measured and calculated pressure ratios.

| | Measure values | | | Calculated values | Percent disagreement |
|---|---|---|---|---|---|
| $\delta_0$ | $p_0$ abs. (p.s.i.) | $p_1$ abs. (p.s.i.) | $(p_0/p_1)$ | $(p_0/p_1)$ | $(p_0/p_1)$ |
| 0.193 | 29.7 | 14.7 | 2.02 | 2.33 | 13.3 |
| 0.270 | 32.7 | 15.7 | 2.08 | 2.15 | 3.3 |
| 0.492 | 34.7 | 19.7 | 1.80 | 2.00 | 10.0 |
| 0.84 | 40.7 | 23.7 | 1.71 | 1.93 | 11.4 |
| 1.36 | 48.7 | 29.7 | 1.66 | 1.73 | 4.0 |
| 2.01 | 54.7 | 34.7 | 1.57 | 1.67 | 6.0 |

This shift in location was necessary because it was not possible to locate a pressure tap precisely at the end of the converging nozzle.

Since the static pressure is very sensitive to area ratio in the vicinity of the throat, the accuracy of the experimental measurement is limited. In Table II the measured and calculated values of $p_0/p_1$ are compared for a series of measurements made with the pressure tap slightly upstream from the throat $(A^*/A_1 = 0.993)$.

Perhaps it should be pointed out that the experimental work, which was directed toward other objectives than the verification of the theory, was terminated before this present analysis was carried out; thus no opportunity to investigate discrepancies has been available.

## BIBLIOGRAPHY


(1) *Jet Propulsion* (Textbook of Air Technical Service Command, Wright Field, 1946), Chapter XV.
(2) J. J. Baranowski and D. A. Seiler, M.S. Thesis in Aeronautics, California Institute of Technology, June 1, 1944.
(3) C. H. Dodge and Leo Zwell, Jet Propulsion Laboratory Report 2-3, April, 1947.
(4) L. L. Foldy, OSRD Report, Section 6.1, No. NS-141, April 25, 1944.
(5) L. Spitzer, Jr., NDRC Report, 6.1—Sr 20—918, July 15, 1942.
(6) P. S. Epstein, NDRC Report, C 4—Sr 30—027, September 1, 1941.
(7) R. C. Brumfield, NOTS Underwater Ordnance Division Report OUM-8.7, April 10, 1946, revised August 8, 1946.

# Fluid Flow Patterns†

GARRETT BIRKHOFF AND THOMAS E. CAYWOOD
*Harvard University, Cambridge, Massachusetts*
(Received February 6, 1948)

A new photographic technique is described, for observing velocity fields in water, induced by the entry of solid missiles. The results of a number of observations are given, and compared with theoretical formulas. The corrections involved in the method are analyzed theoretically, and the theory of the corrections confirmed experimentally. It appears that the method is well adapted to velocity fields 5–100 feet per second, created by objects several inches or more in diameter, but may lead to troublesome corrections at other speeds or with smaller objects.

## I. EXPERIMENTAL TECHNIQUE

### 1. Slichter Multiflash Bubble Technique

THE velocity field of a moving liquid may be observed through time-exposed photographs of illuminated particles[1] suspended in the liquid, or on its surface. This method is often called "chronophotography."

We wish to report on a new technique for observing transient velocity fields in water, with a vertical plane of symmetry, such as occur in water entry and cavity motion problems. In this technique, which is due to Professor L. B. Slichter,[2] *multiflash* pictures are used in place of time exposures; this permits a greater concentration of light energy. Moreover, the particles photographed are located in a vertical plane by the following device.

A thick-walled tube of soft rubber, perhaps 0.5 in. in diameter, is punctured at regular intervals by a wire needle about 0.01 in. in diameter. When the tube is inflated under water at an air pressure of 10–20 lb./in.², air bubbles of diameter $d=0.2$–$2.0$ mm are emitted by each hole. These rise under gravity, in a predetermined, approximately vertical plane. (See Fig. 1.) If a symmetric missile moves through the water in this plane, the bubbles will also move in it, giving a plane section of *three-dimensional motion*.

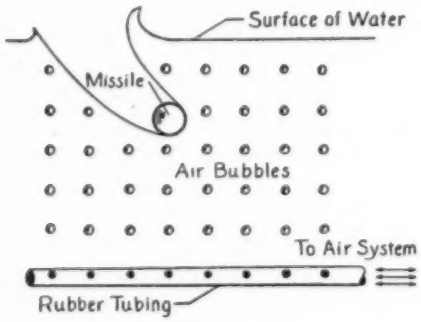


FIG. 1.

If the position of a bubble is located by two light flashes at a time $\Delta t$ apart, and $\Delta x$ is the vector displacement of the bubble as measured directly (by simple proportion) on a photographic plate, then the average vector velocity will be

$$v=\Delta x/\Delta t. \tag{1}$$

A sample photograph is shown on the cover.

The complete experimental set-up is shown in horizontal plan in Fig. 2. When the light source is about 30° to the rear[3] of the bubble plane, the light is reflected from a small spot on the bubble, whose diameter is often as little as one-third of the bubble diameter $d$. The center of the spot can then be located with an error less than $d/12$, so that $\Delta x$ can be measured to about 0.06 mm in favorable cases.[4]

However, we should remember that what we are measuring with 0.06-mm accuracy is the *displacement of bubbles in a finite container*, whereas we wish to infer the corresponding displacement of an infinite extent of water. The "corrections" needed to make this inference are discussed in Sections 9–13.

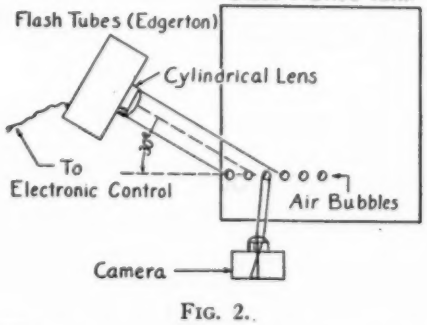


FIG. 2.

† We report here work done under Task B of Contract NOrd-8555 with the Bureau of Ordnance. We wish to thank Professors L. J. Hooper and F. S. Findlayson of the Alden Hydraulic Laboratory of the Worcester Polytechnic Institute, where the experiments were performed, for their advice and cooperation.

[1] See Bibliography at end of this paper: [5], [6], and [7]. For a general discussion of photographic techniques, see [2], pp. 280–94, and [4], Vol. 2, pp. 265–74, and Plates 1–27.

[2] In collaboration with the Morris Dam Group at the California Institute of Technology, working under OSRD contract.

[3] This was also essential for obtaining enough illumination and is incompatible with the use of a plane of light to select a plane of bubbles.

[4] An accuracy of 0.025 mm is claimed in [4], using a prepared emulsion, but the droplets were not confined to a plane.

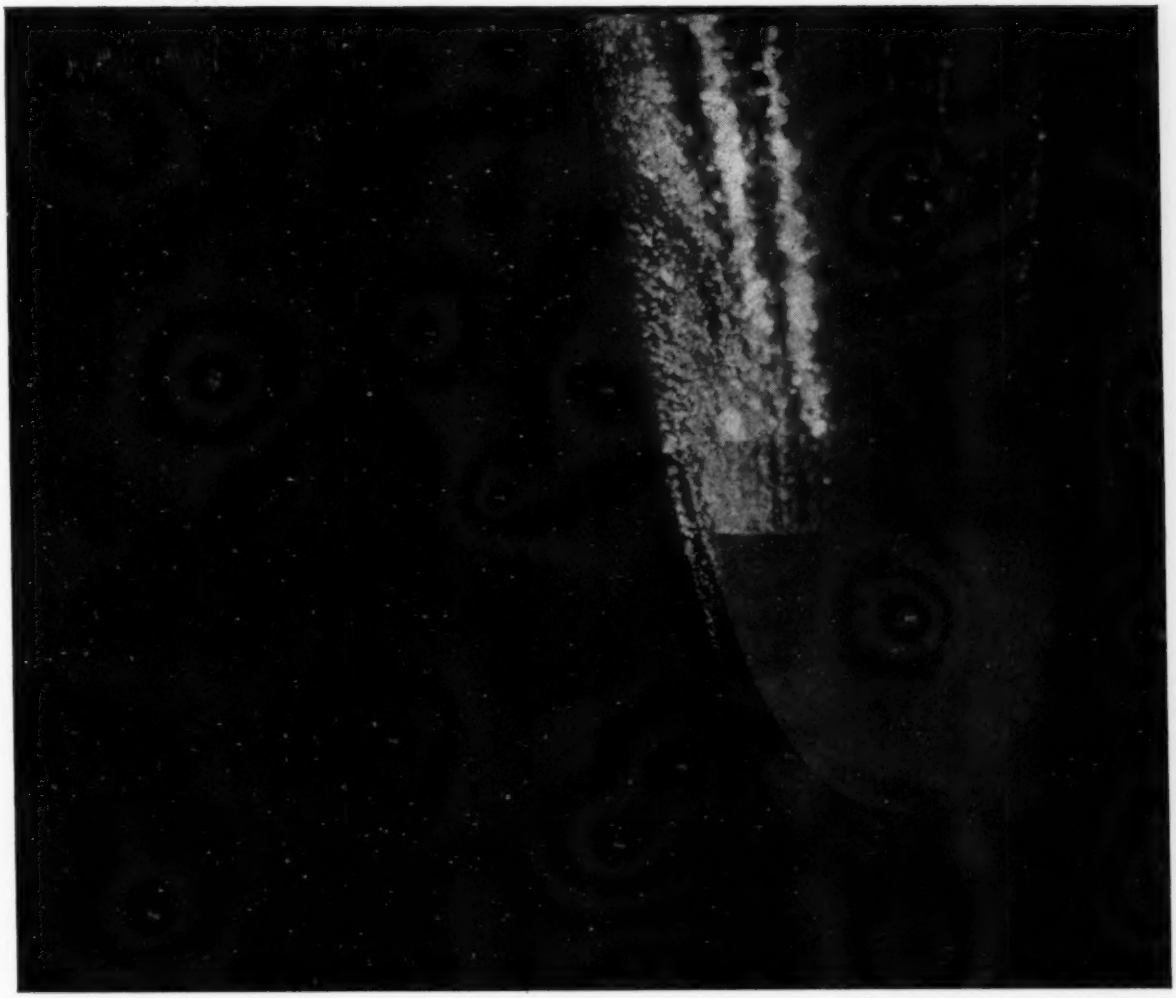

FIG. 3. Vertical cavity due to 8×2-in. prolate spheroid. Model speed 29 ft./sec. at entry. Time interval about 0.003 sec.

### 2. Experimental Difficulties

One of the main experimental difficulties comes in obtaining a uniform array of "good" columns of regularly spaced bubbles. Bubbles more than 0.9 mm in diameter spiral up irregularly, and thus do not stay in a predetermined plane. Bubbles much less than 0.4 mm in diameter rise so slowly as to drift out of the plane with small convection currents.

Most synthetic rubbers we punctured gave irregular individual holes. Commonly, a continuous stream of air would spurt out, and form irregularly into bubbles of various sizes. Each individual hole, carefully punctured in natural soft rubber, gave a "good" column at a suitable air pressure, consisting of bubbles of uniform diameter, spaced perhaps 2 cm apart. At higher air pressures, bubbles would be too large; at lower pressures, they would be too small or not form at all.

The most difficult trick was to make all the bubble columns "good" at the same air pressure. Best results in this regard were obtained by puncturing the tube when wet, inflated with air about 1/5 atmosphere overpressure and stretched about 20 percent. The tube was later used in the same state under about 1-atmosphere overpressure, with the holes about 120° from the top of the tube. By twisting the tube slightly, more uniformity between bubble columns could be achieved—larger bubbles being formed by any hole at angles $\theta > 120°$ with the top of the tube.

Two General Electric FT-2 1 flash tubes with 2000 volts and 9 microfarads were used as light sources, about 30° to the rear of the bubble plane. A separate tube was used for each flash (of the double flash), and the light was concentrated on the bubble plane by individual cylindrical lenses. The flash duration was of the order of $10^{-5}$ second. Since we photographed a field about 24″×18″, we were limited by the total available light intensity. We obtained best results with Super Panchro

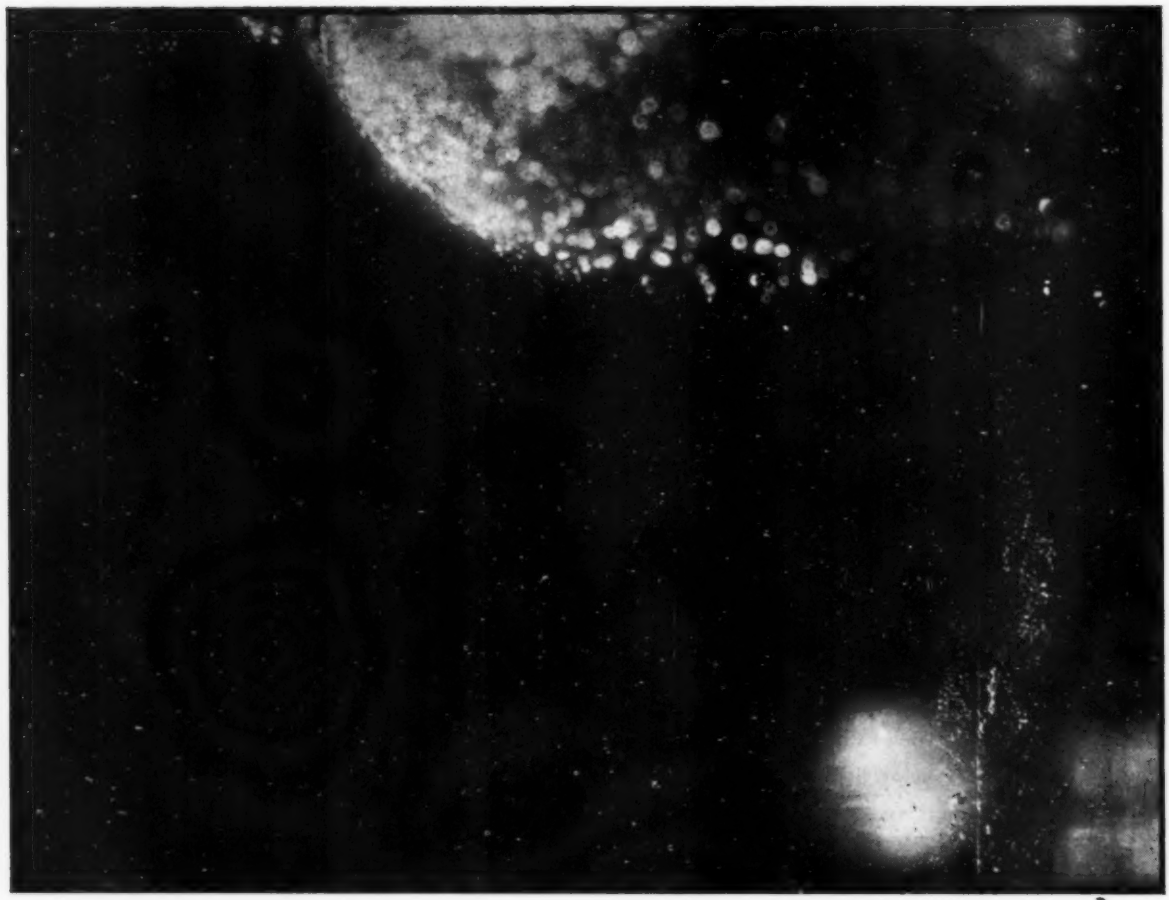

FIG. 4. Deep seal of vertical cavity due to 4×2-in. oblate spheroid. Model speed 29 ft./sec. at entry. Time interval 0.0025 sec.

Press Sports Type film; Super XX film was also satisfactory. A high contrast developer should be used. Best results were obtained with Von-L, and D-72 1:1 was satisfactory.

### 3. Use of Oil Drops

In order to reduce the corrections for virtual mass and "bubble rise" discussed in Part III, we tried replacing air bubbles by droplets of liquids whose specific gravity $\sigma$ was nearer unity. We found (cf. Section 2) that when $\sigma > 0.9$, such droplets rose so slowly as to drift out of the vertical plane. We had best results with hexane, for which $\sigma = 0.69$.

The technique used for air (i.e., punctured rubber tubing under pressure) was not satisfactory for oil droplets, because the oil clogs the holes. When holes large enough to remain open were used, the resulting flow of hexane was too great to be useful. Moreover, when a metal pipe with 0.002-in.–0.02-in. holes was used, the fluid would cling to the sides of the pipe until an enormous bubble was formed. Although some study of fluids, detergents, and metals was made, we were unable to get good results with this technique.

The technique used most successfully was the excretion of hexane from small hypodermic needles. For this purpose #27 hypodermic needles (i.d. 0.0075 in., o.d. 0.016 in.) were used. The needles were mounted vertically on observation tubes which were in turn mounted on a brass tube by means of short links of rubber tubing for shock absorption. An excess pressure of about 50 dynes/cm$^2$ was enough to make bubbles of about 2.4-mm diameter rise 3 cm apart. Less pressure would form the same size bubbles but spaced farther apart. Greater pressure would produce a jet from the needle. This would form into bubbles less than 1 mm in diameter and would be difficult to retain in a plane.

The tip of the needle was left intact. A bubble would form around the opening in the side of the tip, rise vertically to the tip, and then pull free. Thus the effect of the surface tension between the hexane and the metal was greatly reduced.

The needles clogged readily if the pressure was taken off and water allowed to enter the system. The system was cleaned by an internal blast of air and the small wire cleaners that are furnished

with the needles. However, if air got into the hexane line while in operation, the needles would tend to clog.

## II. DISCUSSION OF OBSERVATIONS

### 4. Vertical Entry of Spheroids

We applied the experimental technique described in Sections 1 and 2 to the vertical entry into water of various spheres and spheroids at about 29 ft./sec. entry speed, the unstable vertical orientation being maintained by threading a taut piano wire through the axis of the model.

A typical velocity field is displayed in Fig. 3. It will be noted that the motion of the cavity wall is mainly *radial*, or perpendicular to the cavity axis. This was predicted by Mallock[5] and Bauer. It would also occur in an ideal fluid if the cavity wall were a "free boundary," and the spheroid moving at constant velocity—the velocity being $\psi/2$ forward of the radial direction, where the free boundary makes an angle $\psi$ with the direction of motion. In Plate 1 the boundary layer seems to be perceptible. Just outside of this, the radial component of the fluid velocity seems about one-fourth the velocity $v$ of the spheroid, and the rearward component perhaps $v/10$, near the point of separation of flow.

As the spheroid descends further, the water is accelerated inwards by gravity, so that the cavity collapses on the axis in a "deep seal," described by Mallock. Fig. 4 illustrates the velocity field just after deep seal; a thin jet is rising inside the cavity along its axis.

### 5. Vertical Entry of Wedges

We made studies of a 90° wedge similar to those for spheroids of Section 4. We did this because there was known[6] a complete theory of the steady-state cavity (i.e., wakes) flow past a symmetric 90° wedge in a nonviscous fluid. For a wedge with verticals at (1, 0), (0, 0), and (0, 1), the complex position coordinate $z$ is related to the complex conjugate velocity vector $\zeta$, by the formula

$$1.744z=\frac{4\zeta^3}{(1+\zeta^4)^2}+\frac{\zeta^3}{1+\zeta^4}+\frac{1}{4}\left(W^3Ln\frac{\zeta+W}{\zeta-W}+WLn\frac{\zeta+W^3}{\zeta-W^3}\right)+\frac{\pi}{2\sqrt{2}}, \quad (2)$$

where $W=(1=i)/\sqrt{2}=(-1)^{\frac{1}{4}}$. In Fig. 5 we have

[5] A. Mallock, Proc. Roy. Soc. **A95**, 138–43, (1918); W. Bauer, Ann. d. Physik **82**, No. 7, (1927).

[6] [3], pp. 104–5, formula (27), from which Eq. (2) can be deduced. Flat plates are mathematically simpler, but decelerate too rapidly under free fall to approximate steady-state conditions at all.

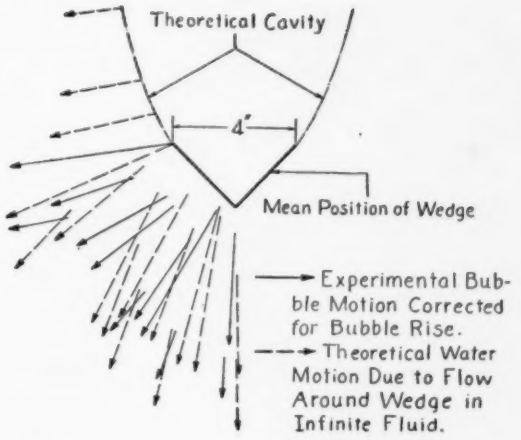
Theoretical Cavity
4"
Mean Position of Wedge
Experimental Bubble Motion Corrected for Bubble Rise.
Theoretical Water Motion Due to Flow Around Wedge in Infinite Fluid.

FIG. 5. Comparison between experimental and theoretical flow around a wedge. Experimental data taken from Plate 35; 4-in.–90° wedge, 2-ft. long. Theoretical flow normalized to produce experimental water motion at bottom of wedge.

compared observed velocities with steady-state theory. The observed velocities are 20 percent–40 percent less than the predicted velocities, and are directed 0°–20° more towards the surface.

We conjecture that both these effects are due to the presence of the horizontal free surface, but the second may be due to the correction of Section 14.

### 6. Two-Dimensional Jet

The two-dimensional cavity behind a 4-in. diam. wedge entering at 29 ft./sec., collapses in a "deep seal," much like the axially symmetric cavity behind a spheroid (cf. Section 4). We display in Figs. 6 and 7 (see cover) both the pinching off of the cavity before deep seal, and the approximate velocity field of the up-jet, which is formed after deep seal above the point of deep seal. Thus, in Fig. 6, the water is moving inwards towards the cavity (which appears light gray), and also downward in

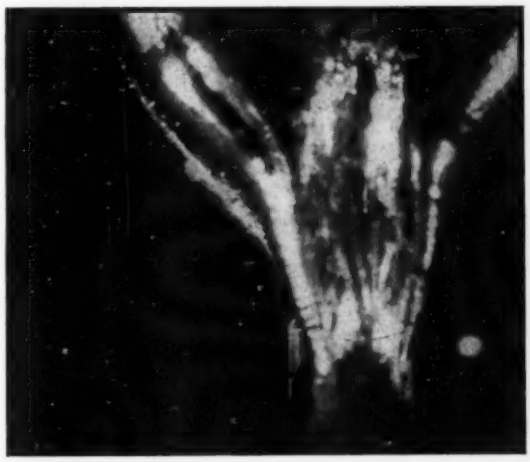

FIG. 6. Deep seal of 90° 2-ft.×4-in. wedge. Model speed 25 ft./sec. at entry. Time interval 0.0025 sec.



UNIVERSITY OF MICHIGAN LIBRARIES

FIG. 7. Deep seal of 90° 2-ft.×4-in. wedge. Model speed 25 ft./sec. at entry. Time interval 0.0025 sec. Shows flow accompanying up-jet formed at surface.

the upper portion of the picture. Other photographs reveal the velocity field associated with the down-jet formed below the point of deep seal.

We were able to make a mathematical model of an infinite two-dimensional jet with a free surface as follows. Suppose the velocity $\bar{\zeta}$ at infinity is 1, we have a physical $z$-plane, a complex velocity potential $W=U+iV$, and a conjugate velocity $\zeta=dW/dz$. One can map a complex half $t$-plane onto the shaded region of the $W$-plane (see Fig. 8) by the generalized Schwarz-Christoffel transformation

$$dW=\zeta dz=(tdt/t+1). \tag{3}$$

One can map the complex half of the $\tau$-plane onto the shaded area of the $\zeta$-plane (hodograph plane) by setting $2\tau=\zeta+1/\zeta$. Then $t(0)=0$, $\tau(0)=\infty$; $t(B)=-1$, $\tau(b)=1$; $t(A)=\infty$, $\tau(A)=-1$. These points correspond under the (unique) linear fractional Moebius transformation, $t=-2/(\tau+1)$. Substituting, we get $t+1=(\zeta-1)^2/(\zeta+1)^2$, $dt=4(\zeta-1)/(\zeta+1)^3$, $tdt/(t+1)=-16\zeta/(\zeta-1)(\zeta+1)^3$, and so

$$dz=-16/(\zeta-1)(\zeta+1)^3. \tag{4}$$

This can be integrated in closed form, as

$$z=-\frac{4}{(\zeta+1)^2}-\frac{4}{\zeta+1}+2\ln\frac{\zeta+1}{\zeta-1}+C. \tag{5}$$

In Fig. 9 we have compared the flow pattern (20) with Fig. 7; the agreement is surprisingly good, in view of the facts that gravity is not negligible, that the flow is not stationary, and that we have a horizontal free surface.

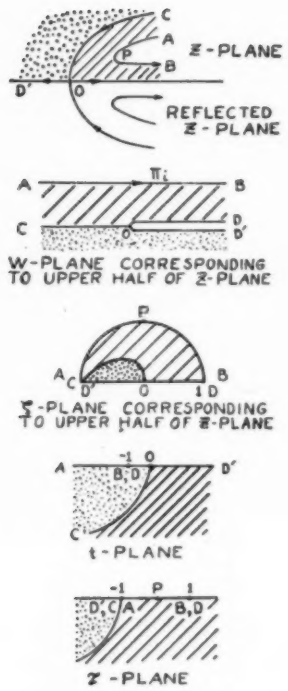


FIG. 8.

## 7. Oblique Entry of Spheres

We also studied the velocity fields associated with 1.5–3.0-in. diam. spheres entering water obliquely, at speeds of 80–150 ft./sec.

The instantaneous streamlines in the vertical plane of motion are approximately concentric circles, centered in the point of impact.

In Fig. 10 we have compared the vector bubble displacements, as observed in one of our photographs, with the field of a uniform line of sources, and reflected line of sinks, as sketched in Fig. 11. (The graphical construction of this field is greatly facilitated by use of Ex. 3 on p. 10 of Kellogg's *Potential Theory*.)

It will be noted that the velocity field is far from being "axially symmetric" about the axis of motion—greater above the cavity than below it.

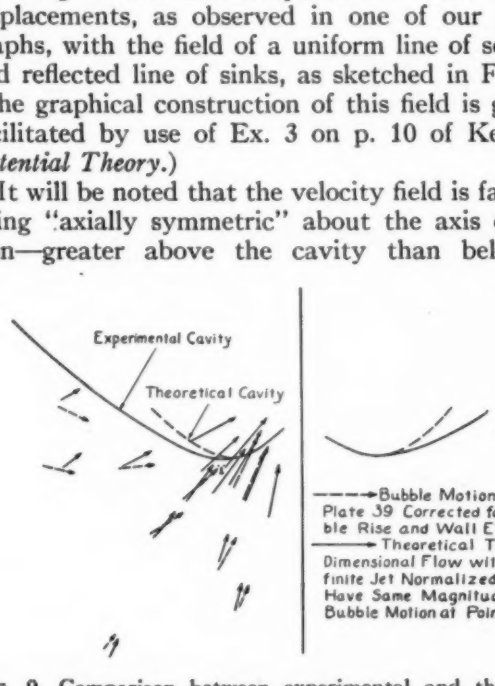
Experimental Cavity
Theoretical Cavity
Bubble Motion from Plate 39 Corrected for Bubble Rise and Wall Effect.
Theoretical Two-Dimensional Flow with Infinite Jet Normalized to Have Same Magnitude as Bubble Motion at Point "1".

FIG. 9. Comparison between experimental and theoretical $z$-dimensional jet flow.

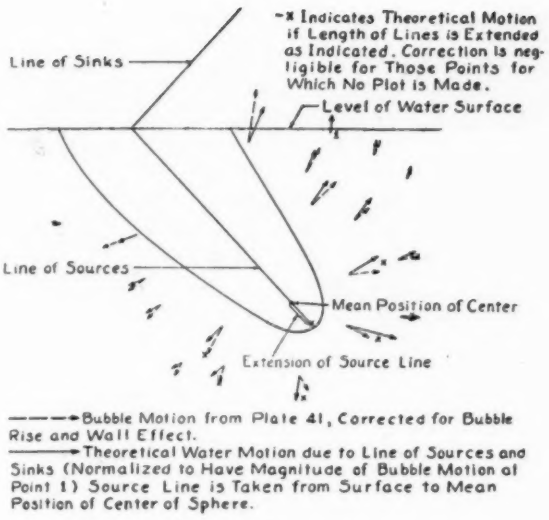
-x Indicates Theoretical Motion if Length of Lines is Extended as Indicated. Correction is negligible for Those Points for Which No Plot is Made.
Line of Sinks
Level of Water Surface
Line of Sources
Mean Position of Center
Extension of Source Line
——→Bubble Motion from Plate 41, Corrected for Bubble Rise and Wall Effect.
——→Theoretical Water Motion due to Line of Sources and Sinks (Normalized to Have Magnitude of Bubble Motion at Point 1) Source Line is Taken from Surface to Mean Position of Center of Sphere.

FIG. 10.

This asymmetry of the velocity field is consistent with the observed fact[7] that the surface effect on hydrodynamical forces is small, if a sphere is four diameters below the surface. A sink $n$ diameters above the surface would cause velocity fields at points $n-\frac{1}{4}$ and $n+\frac{1}{4}$ diameters below the surface, differing only in the ratio $1\pm\frac{1}{8}n^2$. If we consider that most of the velocity is due to sources below the surface and compute pressure variation from Bernoulli's equation (3) with $\partial U/\partial t=0$, we are led to expect an unbalanced force of at most 1 percent at four sphere diameters submergence.

### 8. Surface Seal and Ricochet

It is known[8] that at high Froude numbers $F=v^2/gd>5000$ or so, the "deep seal" of Section 4 is replaced by a "surface seal" which may be described somewhat as follows. The splash is blown inward by air rushing into the cavity. This forms a constriction or "neck." The air rushing through this neck, causing a local underpressure, is $\frac{1}{2}\rho_1 v^2$, where $\rho_1$ is the air density, and $v$ is at least the missile velocity. This airflow probably separates from the cavity wall at the edge of the splash, leaving a turbulent air-filled or foamy region at mixed pressure between the inrushing air and the cavity wall. Some water is thrown against the cavity wall, which it makes turbulent and opaque. The underpressure causes the neck to contract more and more, choking off air from the cavity, until a complete "surface seal" is formed.

[7] This was observed by C. Ramsauer, *Uber den Ricochetschuss*, Thesis (Kiel, 1920). The rough computation below seems to us clearer than Ramsauer's "pressure zone" explanation.

[8] See [11], [12]; also Mallock (reference 5). Roughly, surface seal supersedes deep seal when $\frac{1}{2}\rho_1 v^2$ exceeds the hydrostatic overpressure $3\rho gd$ at the depth (3 diameters) of deep seal—or when $F>6\rho/\rho_1=4400$, roughly.

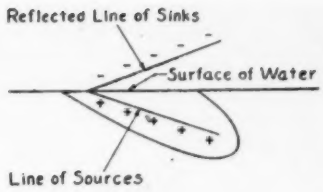


FIG. 11.

We were fortunate in photographing the approximate velocity field in the neck of the cavity. In Fig. 12, the rear side of the neck is seen to be moving inwards horizontally; this we found to be true in all cases. The front (upper) side of the cavity, on the other hand, usually moved upwards at first, but later reversed its direction of motion, at least with sufficiently oblique entry. The "dipole" field of secondary flow, clearly revealed in Fig. 13, is just beginning to form under the cavity.

More recent photographs show that, in the case of shallow entry angles (35° or less with the horizontal), after surface seal is complete, a dipole-like velocity field travels vertically downward, away from the cavity. This is illustrated in Fig. 13, and is presumably due to the *down-jet* accompanying surface seal. This down-jet is especially visible in photographs taken at steep entry angles, and can be distinguished in these cases from the indentations of the cavity wall caused by inward-blown splash.

Another interesting series of pictures concerns the *ricochet* of a 3″ aluminum sphere at 7° impact angle (see Fig. 14). It is remarkable that the water velocity tends to be *tangential* to the cavity at the end of the long plough, and not radial, as in most cavity motion. It is true that this is necessary if the circulation around the closed curve of Fig. 15 is zero (as required by Kelvin's theorem for nonviscous fluids), if the flow along the free surface is

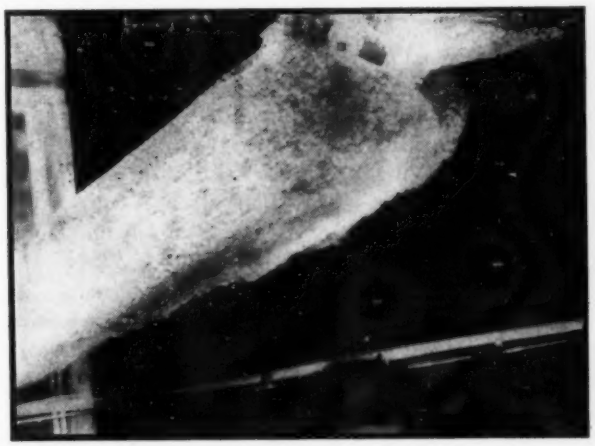

FIG. 12. Surface seal behind 1.5-in. diam. sphere, which entered at 150 ft./sec., with trajectory 30° below horizontal. Time interval 0.004 sec.

FIG. 13. "Dipole" moving down from point of surface seal, behind sphere which entered at 250 ft./sec., with trajectory 15° below horizontal direction. Time interval 0.004 sec.

mostly normal to the surface, and if the interior streamlines escape to the free surface, as they do. In Figs. 14 and 15, the missile has moved out of the photographic field to the right.

## III. CORRECTIONS

### 9. General Remarks

The observations described above are subject to three corrections.

The first is the "wall effect" due to the finite size of the container. We estimated it on the basis of the Dirichlet flow around a 4-in. diameter sphere in a 32-in. diameter spherical tank, as given in [3], Section 93, pp. 124–5. We did this because we worked at least 16 in. from the wall of a 52×52-in., 34-in. deep tank. The estimated effect seemed small.

The other two involve the relative motion of the bubble in the water, due to the difference between bubble density and water density. This difference causes the bubbles to rise under their own buoyancy. The rate of rise $v_b$ of bubbles in each column can be measured; it is normally between 3 and 8 cm/sec., and consistent to within ±10 percent in each "good" column. We simply subtracted a vertical "bubble rise correction" of $v_b\Delta t$ from each observed displacement. This was usually small (cf. the bubbles near the edge of each photograph); however, because of it, *the multiflash bubble technique for observing velocity fields is only useful for fluid velocities above* 30 cm/sec.

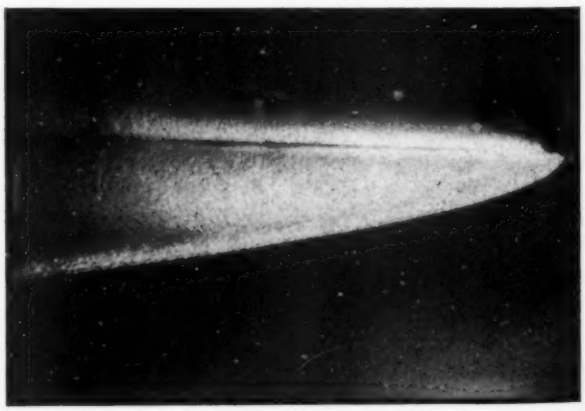

FIG. 14. Ricochet of 3-in. aluminum sphere. Entry angle 8°, velocity 115 ft./sec. Time interval 0.0018 sec. Sphere has left the picture on the right.

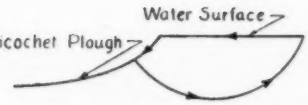


FIG. 15.

The difference between bubble density and water density also makes the bubbles respond more violently to pressure gradients (i.e., accelerations) than the surrounding water. The resulting "virtual mass" correction is the main difficulty with the multiflash bubble technique; we shall now discuss it directly. It should be observed that the problem has arisen before, in connection with the theory of suspensions.[9] However, the treatment sketched below, and first given in [13], offers some novelty.

### 10. "Virtual Mass" Correction

We estimated the "virtual mass" correction, which might be expected in the absence of gravity (buoyancy) and walls, by assuming equilibrium between the following three forces on each bubble.

Force 1. A pressure thrust $m\mathbf{a}$, where $m$ was the mass of water displaced by the bubble, and $\mathbf{a}$ was the acceleration of the surrounding water.

Force 2. An inertia "force" of $m\sigma\mathbf{a}_1+\frac{1}{2}m(\mathbf{a}_1-\mathbf{a})$, where $\sigma$ was the specific gravity of the bubble, and $\mathbf{a}_1$ was the bubble acceleration. Here $m\sigma$ is ordinary inertia, and $\frac{1}{2}m$ is the "induced mass" inertia. ([3], p. 124).

Force 3. A resistance $D$ to the relative bubble velocity $\mathbf{v}_1-\mathbf{v}$($\mathbf{v}_1$ being bubble velocity, and $\mathbf{v}$ the velocity of the nearby water). We estimated this from empirical data, for *steady* relative motion. A reasonably good fit (cf. Fig. 16) for an undistorted air bubble is given by

$$C_D=0.4+24/R, \tag{6}$$

for $R<10^5$. Here $R=|v-v_1|d/\nu$ is the Reynolds number. For distorted bubbles, we used (7) (below) and

$$C_D=0.4\Delta+24/R. \tag{6'}$$

$\Delta$ is the ratio of actual cross-section area to undistorted cross-section area.

There is no theoretical justification for assuming that Forces 1–3 are independent, but we could think of no better assumption. If the assumption is made, the condition of equilibrium becomes

$$m(\sigma+\tfrac{1}{2})\mathbf{a}_1-1.5m=D\mathbf{a}=\tfrac{1}{2}\rho(v_1-v)^2AC_D. \tag{7}$$

where $A=\pi r^2$ is the cross-section area of the bubble, and $C_D$ is the drag coefficient.

### 11. Case of Impulsive Velocity: Bubbles

In case the water is given an *impulsive velocity* $\mathbf{v}$, which is thereafter held constant, it is easy to deduce from Section 10 that the initial relative velocity $\mathbf{v}_1-\mathbf{v}$ of an air bubble should be $V=2\mathbf{v}$, as in Section 3. This relative velocity should then die out according to the differential equation $dV/dt=\alpha V^2+\beta V$, where the constants $\alpha$, $\beta$ are determined by formulas (7)–(6') with $\mathbf{a}=0$ in (7). This is easily integrated, giving a predicted ratio $s_1/s$ of bubble displacement $s_1$ to water displacement $s$, for each time interval $\Delta t$ and sphere density $\sigma$. Two curves obtained in this way are shown in Figs. 17–18.

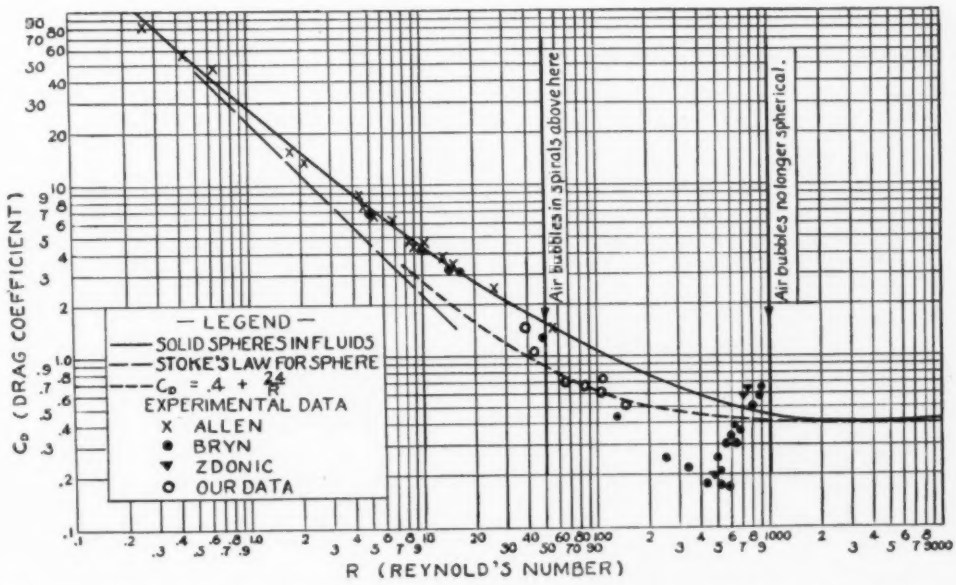


FIG. 16.

[9] A. B. Bassett, *A Treatise of Hydrodynamics* (Cambridge University Press, London, 1888), Vol. 2, Chapter V; C. W. Oseen, *Hydrodynamik* (Leipzig, 1927), p. 132; T. Chan-Mou, Mededelung 51 Lab. Aer. Hysr. Delft; G. I. Taylor and M. Glauert, Reps, and Mem. 2025 Aer. Res. Comm; E. Brun and M. M. Vasseur, Belleveue No. 3 (1947); Irving Langmuir and K. B. Blodgett, Office Tech. Services, U. S. Dept. Commerce, No. 27, 565.

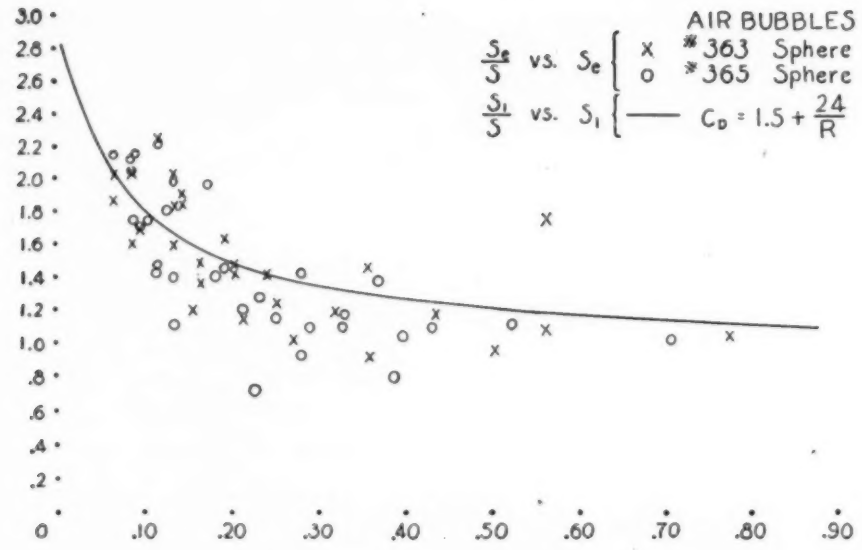


FIG. 17.

We checked this prediction experimentally as follows. Suppose a half spheroid (ellipsoid of revolution), is floating with its equatorial plane on the surface of the water. Let the hemispheroid be given a downward impulsive velocity.[10]

Since the pressure on the water surface is constant, the pressure gradient will be normal to it. Hence, so will the acceleration. Integrating, the impulsive velocity imparted to the hemisphere will also be vertical; this was also confirmed experimentally by our photographs. We conclude that the velocity potential $U$ satisfies

$$U = U_0 = \text{const. on the } (x, z)\text{-plane initially.} \quad (8)$$

But by the well-known *principle of reflection*, this determines mathematically a unique solution of (1), given zero velocity at infinity, and the normal velocity $\partial U/\partial n$ on the boundary. This solution corresponds (cf. reference 9) to supplementing the half spheroid by its mirror image in the water surface—i.e., by taking the corresponding flow induced by a spheroid, relative to fluid at rest at infinity. In the case of ideal fluids, this flow is well known ([3], p. 141), and reduces to a *dipole* field in the special case of a hemisphere.

In this special case, the theoretical angle $\theta_t$ between the vertical and the direction of fluid motion at a point $P$, subtending a polar angle $\phi$ (see Fig. 19 where 0 is the mean sphere center), satisfies $\theta_t = \phi + \tan^{-1}\frac{1}{2}\tan\phi$. In Fig. 20 we have plotted as a function of $\phi$, the difference $\theta_e - \theta_t$ between $\theta_t$ and the experimental flow direction $\theta_e$, corrected for bubble rise and wall effect. This is

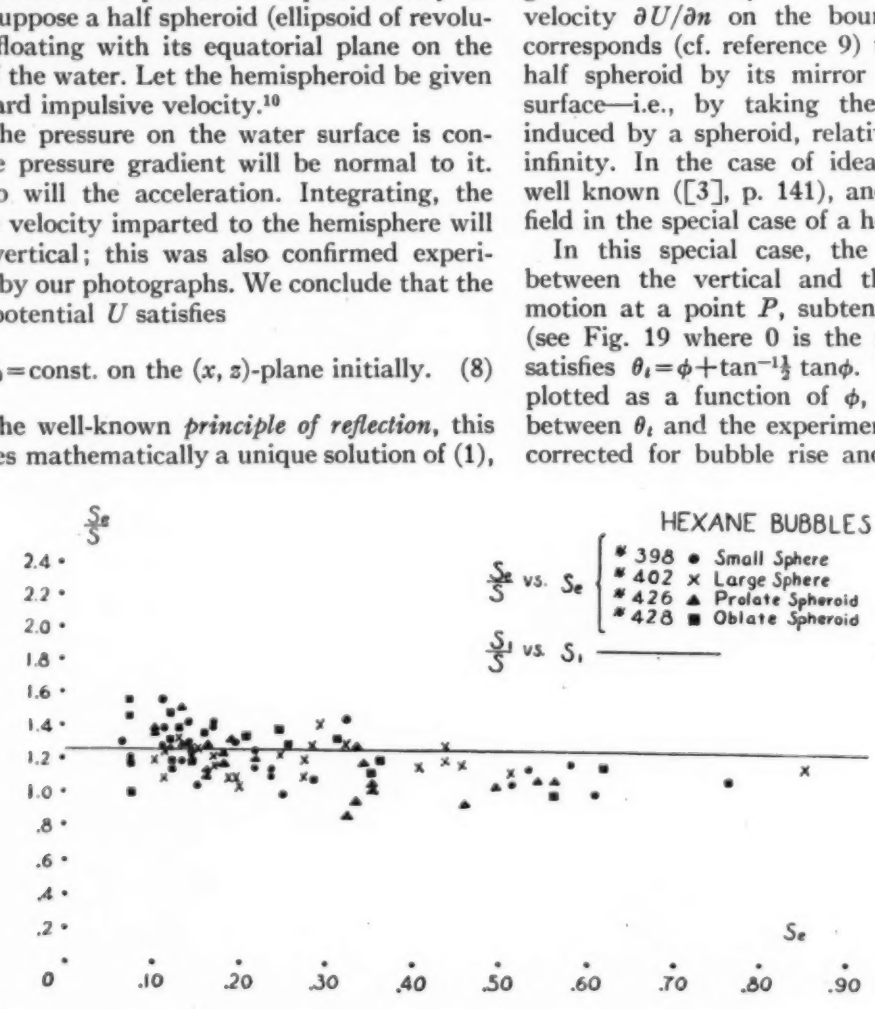


FIG. 18.

[10] This was accomplished by threading the model on a 0.040″ vertical steel wire under tension and dropping a 10-lb. weight down the wire. The acceleration of the model took less than 0.0002 sec. and its velocity was about 25 ft./sec. We had trouble with wire rupture and horizontal vibration until we stabilized the weight and placed a 1/16-in. shock absorber below the weight.

one of the first[11] quantitative comparisons of predicted velocity fields with photographed velocity fields. It will be seen that there is a scatter of $\pm 5°$, but that a definite disagreement is present for observed points near the surface ($\phi$ nearly 90°). This may be due to the finite displacement. If we remember that the actual water surface is above the mean sphere center, and that theoretically, at all times (see Section 12), $\theta = 180°$ at the surface, we are led to plot also $\theta_c - \theta_t^*$ against $\phi$, where $\theta_t^*$ is the angle between the vertical and the velocity vector as $P$ for flow around a sphere with center at the surface.

In view of Sections 3–4, the ratio $s_c/s$ of the photographed bubble displacement $s_c$ (corrected for bubble rise) to the theoretical water displacement $s$, should vary between one and three. In Fig. 17, we have compared the theoretical curves described above, with observed values of $s_c/s$, plotted against $s_c$. Note the data tends to lie above the theoretical curve for $s_c < 0.15$ and below for $s_c > 0.15$. This may be explained by the fact that the distortion of the air bubbles is not constant but greater for areas which are nearer the spheres, and thus have greater motion.

On the other hand, the evident general agreement between the predicted curve and the observed points confirmed our view regarding the nature and magnitude of the "virtual mass" correction.

### 12. Impulsive Velocity: Check with Oil Drops

Analogous experiments were made using hexane bubbles. In Fig. 20b, $\theta_c - \theta_t$ is plotted against $\phi$ for four spheroids. There is a dispersion of about $\pm 4°$ in most regions, presumably due to errors in observation. As in the case of air bubbles (Section 7) one notes a discrepancy between the observed and theoretical data near the surface, although below the surface ($\phi < 90°$) this is not so noticeable.

As in Fig. 17, we plot in Fig. 18 the resulting "theoretical" curves of $s_1/s$ and the experimentally observed ratios $s_c/s$ for the case of the four spheroids with hexane bubbles. Here we took $\Delta$ in (6′) to be 2.5 (i.e., the distorted cross section was 2.5 times the undistorted area). Again the scatter is of the order of $\pm 10$ percent, but the general agreement between the predicted curve and the observed point is apparent. As in Fig. 17, one notes that for high values of $s_c$, the data lie below the theoretical curve. This may be explained as before in terms of the distortion of the bubbles.

We note that in the case of hexane, the ratio of $s_1/s$ never exceeds 1.26. If the specific gravity of the "improved" fluid were 0.85, this ratio would never exceed[12] 1.1.

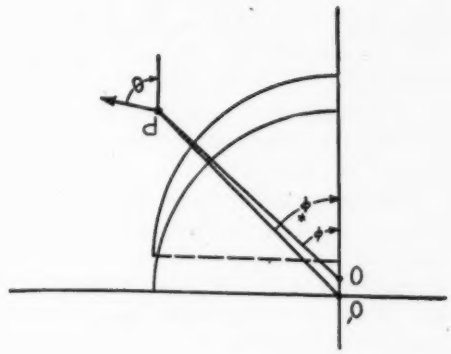


FIG. 19.

The case of a disk entering water should be the limiting case of an oblate spheroid. Figure 21 is a plot of $\theta_c - \theta_t$ against the polar angle $\phi$ from the mean center of the disk, the motion of the disk being less than $0.1d$. The discrepancy near $\phi = 90°$ is more noticeable here than in any other case (cf. Figs. 20a–20b). This is probably because the water pushed out radially prevents the motion on the surface near the disk from being vertical.

A comparison of the ratio of the magnitudes is made in Fig. 22. It is noted that in this one case, all the data lie above the theoretical curve and in fact, whereas the theoretical curve has as its average value about 1.24, the experimental data have an average value of 1.45. We are not able to explain this. Though the experimental errors were larger with oil drops than with air bubbles, we believe they were insufficient to explain the preceding discrepancy.

### 13. Case of Impact

If an object with principal radii of curvature $r$, $r'$ at the point of impact is dropped into still water with speed $v$, the rate at which water volume is displaced per unit time (i.e., the cross-section area at the surface level) is given at first by $2(rr')^{\frac{1}{2}}v^2t$, plus terms of higher order in time $t$ after impact. It seems reasonable to suppose (cf. Section 4) that at large distances, the motion of the water corresponds to a dipole with proportional strength. Therefore at a fixed point of the water, the instantaneous velocity should be fixed in direction, but have magnitude proportional to $t$—i.e., *the vector acceleration* **a** *should be approximately constant.* In this case, one can integrate (7) to get

$$\frac{t}{\sigma + \frac{1}{2}} = \frac{1}{A} \log \left[ \frac{2k_2v_2 + k_1 + A}{2k_2v_2 + k_1 - A} \cdot \frac{k_1 - A}{k_1 + A} \right], \qquad (9)$$

[11] None is given in [5], [6], or [7]. We have not found one of Prandtl, Plate 1(a). A comparison of Pitot tube observations with theory has been made by Bryant and Williams, *Phil. Trans.* **225A**, 199–245 (1925).

[12] The factor 3 for air has the unsuspected advantage that it magnifies motions far from the missile which might otherwise be difficult to observe.

where $A=[k_1^2+4(1-\sigma)ak_2]^{\frac{1}{2}}$, $k_1=18\nu d^2$, $k_2=3/10d$, $a$ is the magnitude of **a** and $d$ is the diameter of the bubble.

By means of (9), we may plot the bubble velocity against time for each value of $a$. These curves are shown in Fig. 23.

We compared the results of the preceding calculations with multiflash pictures of the velocity field of a sphere entering water vertically, previously taken by Professor Slichter. The comparison of theory with experiment is shown in Fig. 23, points associated with the same bubble being connected with the same value of $a$. Again we have good qualitative agreement.

It is to be noted that if there were no relative bubble motion, the curves (9) of Fig. 23 would be replaced by straight lines. The large discrepancy brings out the fact, not previously realized, that *the multiflash bubble technique is unsuitable for studying velocity magnitudes just after impact.*

### 14. Circular motion and general conclusions

In case the water moves in a *circle* of radius $R$, at constant angular velocity $W$, the correction

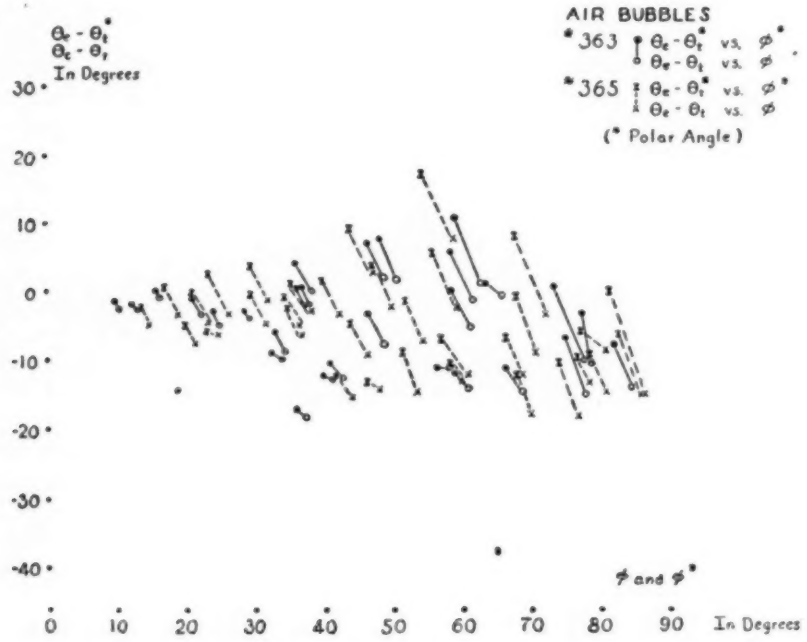


FIG. 20a.

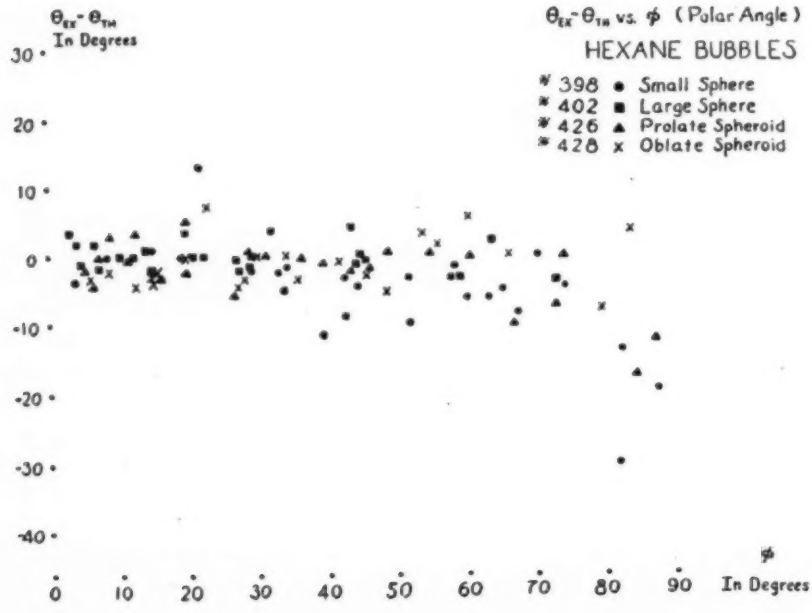


FIG. 20b.

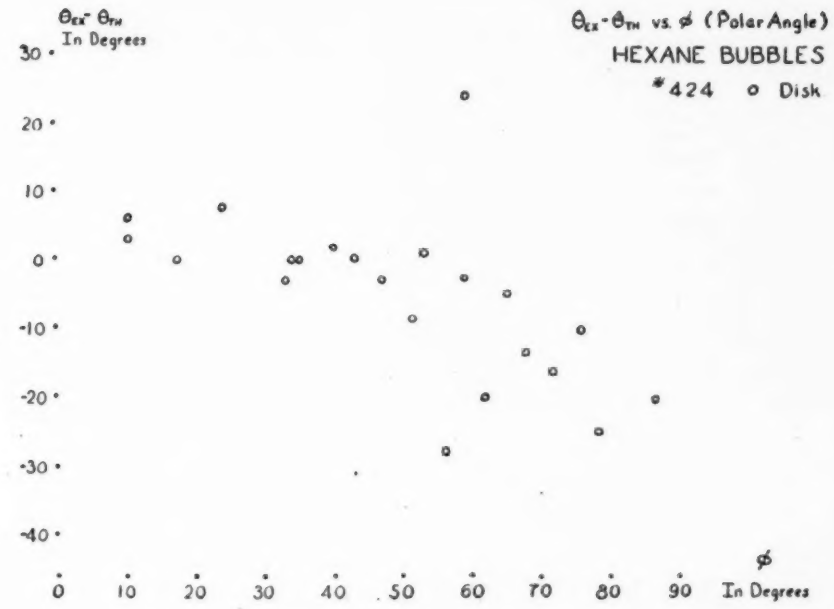


FIG. 21.

predicted by (6′)–(7) can again be found. The bubble will move in another circle, also with angular velocity $W$, so that $a_1-a$ will be 90° out of phase with $v_1-v$, and will be $W$ times as great. Hence, Forces 1, 3 of Section 10 will form the sides of a right triangle, whose hypotenuse will be Force 2. To a first approximation, this leads to the prediction that *in the case of circular motion, the phase of the bubble velocity will be* $(8r/3C_DR)^{\frac{1}{2}}$ *radians ahead* of the water velocity vector. Figure 24 gives contours of this phase advance (in degrees) against $r/R$, for various values of $C_D$.

In the case of vertical entry of 1.5-in. diam. spheres at 150 ft./sec., this predicted phase advance can attain a maximum value as high as 40° or more.[13] These maximum values are attained for bubbles at some distance from the cavity, and therefore having orbits whose radii are small.

The predicted phase advances can, in some cases, be checked against observation. On many double-flash pictures, two cavity outlines are superimposed. Where the outlines cross, the water displacement is tangential. The measured angle between this tangential direction, and the observed direction of bubble displacement, agreed very well in all cases with the predicted phase advance. This agreement confirms further our correction formulas.

The predicted "virtual mass" corrections are considerably smaller for 4-in. diam. missiles than for those of 1.5-in. diam.—perhaps of the order of 10–20 percent of the velocity near the cavity and in front of the missile, and 15–35 percent further to the side. In general, pictures taken by the multi-flash bubble technique tend to exaggerate motion in the direction of acceleration. If $C_D \geqq 0.4$, 99 percent of the relative velocity is lost in eight (bubble)

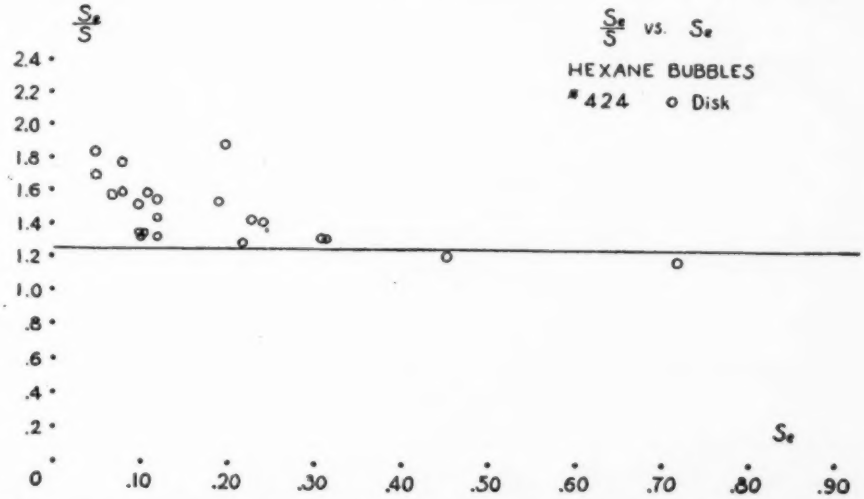


FIG. 22.

[13] The measurements and computations underlying these statements were made by Mr. Robert Isaacs.

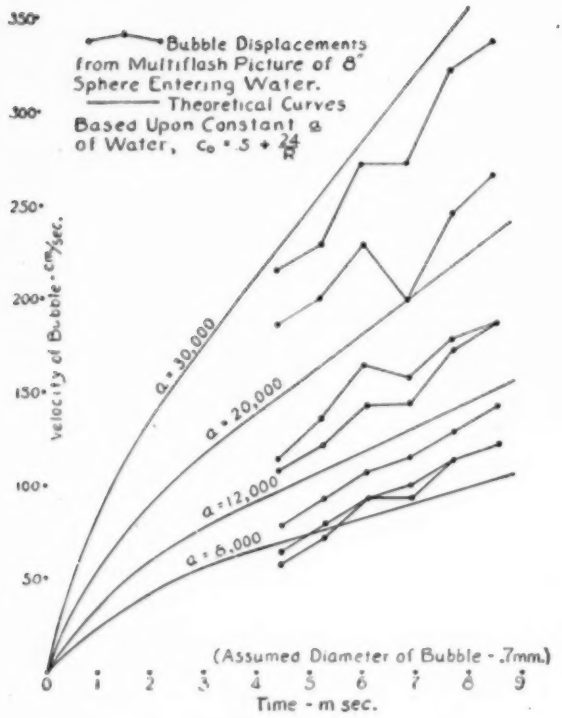
Bubble Displacements from Multiflash Picture of 8" Sphere Entering Water.
Theoretical Curves Based Upon Constant a of Water, $C_D = .5 + \frac{24}{R}$
Velocity of Bubble - cm/sec.
a = 30,000
a = 20,000
a = 12,000
a = 6,000
(Assumed Diameter of Bubble - .7mm.)
Time - m sec.

FIG. 23.

diameters of bubble motion, if the water undergoes no acceleration in this time. Hence, in practice, the relative displacement is limited absolutely by a few millimeters, except insofar as the instantaneous acceleration a cm/sec.[2] of the water tends to produce a relative velocity of the order of $(v_1-v) \simeq (a/2)^{\frac{1}{2}}$ cm/sec.

## APPENDIX

During the course of our work, we made some observations on the behavior of bubbles, which may be worth recording.

Our measurements of the rate of rise of air bubbles in water, in the region $40<R<200$, gave values of the $C_D$ which were considerably lower than might be inferred from the standard references (*Handbuch. Exp. Physik*, Vol. 4, Part II). We discovered that a similar disagreement was found in recent independent measurements by Bryn and Zdonik,[14] and have plotted the data in Fig. 16. (For $R>50$, the bubbles do not rise in a straight line, but spiral upward; for $R>1000$, they are no longer spherical.)

However, these data must be examined rather critically. There is a large mutual influence of two spheres. At very low $R$, Stimson and Jeffery (Proc. Roy. Soc. **A111**, 110 (1926)), have deduced a correction factor $3d/4D$, where $d$ is the bubble diameter, and $D$ the distance apart. This is corroborated visually by watching the apparent "attraction" of bubbles in the same column, and has been estimated empirically as 10 percent by Mr. Jackson of H.M.U.D.E., Portland.

We also investigated the ability of bubbles to maintain their spherical shape in a pressure field. In order to have the change in hydrostatic pressure through one bubble diameter $d$ be less than the difference between the pressure inside and outside the bubble, caused by surface tension, we must have $d^2<30$ cm; with hexane drops, we get the same result.

Under an acceleration of $Ng$, the corresponding figure would

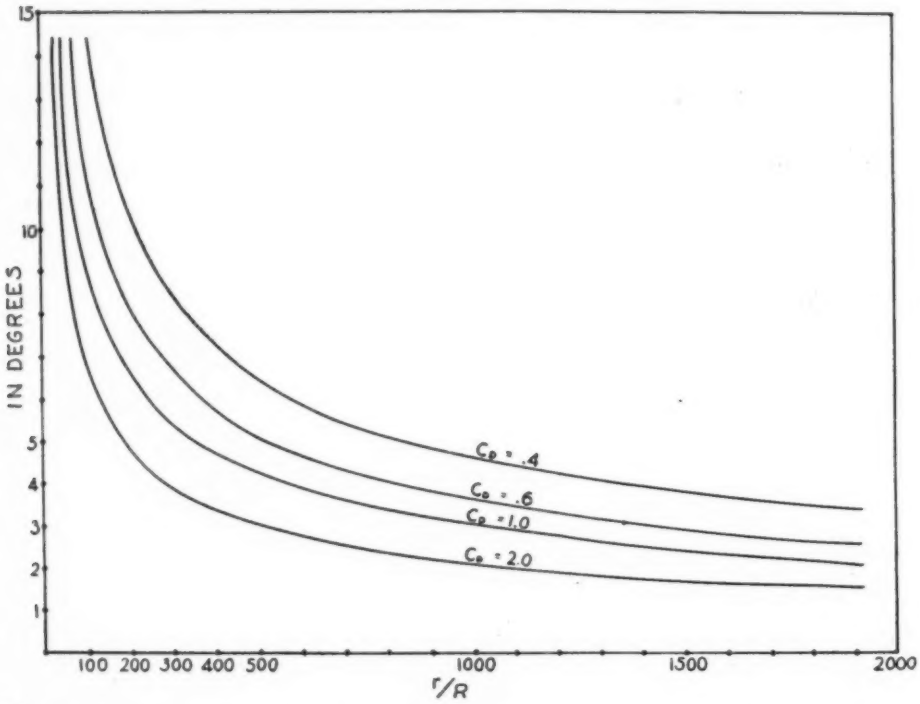


FIG. 24.

[14] T. Bryn, Forschung **4**, 27 (1933); Zdonik, *Absorption of Gases in Liquids*, M.S. Thesis, Chem. Eng. M.I.T. (1942). See also O. Miyagi, Tech. Reps. Tohoku Imp. Univ. (Sendai) **8**, 587–600 (1929), and references given there. We measured the rate of rise by a double flash exposure with a disk rotating at known speed. We estimated the bubble diameter by volumetric measurements, by silhouette pictures, and by measuring the distance between two reflecting spots on the bubble, assumed to be spherical.

be $Nd^3 < 30$ cm—hence, if $d = 0.7$ mm, bubbles should be able to support accelerations up to the order of 6000g, but not more. This may explain why we found it difficult to get good pictures of bubbles in the region immediately in front of spheres moving at 150 ft./sec. Of course, when bubbles are distorted, there is a small correction to be made to account for the shift in position of the spot on the bubble from which light is reflected.

Silhouette pictures like Fig. 25 show the bubble distortion directly; in such cases, we measured the motion of an estimated centroid of the bubble. Fig. 25 is of oildrops under impulsive motion of a spheroid (Section 11); the bubbles nearest the spheroid were dished out at the center, and no longer even convex. Under similar circumstances, air bubbles simply exploded, appearing like bunches of grapes in silhouette photographs. These observations place a further limitation on the accelerations to which the multiflash bubble technique can be applied. On the other hand, empirical evidence indicated that this limitation was less severe than the rough calculations outlined above had led us to suppose.

FIG. 25. Silhouette of 4×8-in. prolate hemispheroid with hexane bubbles. Presubmerged model struck by 5-lb. weight traveling 29 ft./sec. Time interval 0.0017 sec. Note distortion of hexane bubbles due to impact.

## BIBLIOGRAPHY


[1] C. Cranz, *Lehurbuch der Ballistik* (B. G. Teubner, Leipzig, Germany, 1910–31).

[2] S. Goldstein, *Modern Fluid Dynamics* (Oxford University Press, London, 1938), 2 Vols.

[3] Horace Lamb, *Hydrodynamics* (Cambridge University Press, London, 1932), sixth edition.

[4] L. Prandtl and O. S. Tietjens, *Hydro- and Aeromechanics* (New York), 2 Vols.

[5] G. Marey, Comptes Rendus **116**, 913–24 (1893).

[6] Jones, Farren, and Lockyer, Reports and Memoranda of Aer. Res. Comm. **1065**, 54–65 (1926–27).

[7] P. B. Walker, Reports and Memoranda of Aer. Res. Comm. **1402**, 97–171 (1931–2).

[8] A. E. Worthington, *A Study of Splash* (Longmans Green and Company, New York, 1908).

[9] Th. von Kármán, "Impact of seaplane floats," NACA Tech. Note 321.

[10] H. Wagner, "Uber die Entstehung des dynamischen Auftriebes von Tragflügeln," Zeits. f. angew. Math. Mech. **5**, 26 (1925).

[11] R. M. Davies, "The influence of atmospheric pressure on . . . the fall of small scale projectiles into a liquid" (1944) (unclassified mimeographed report).

[12] D. Gilbarg and R. A. Anderson, "Influence of atmospheric pressure on . . . entry of spheres into water," J. App. Phys. **19**, 157–169 (1948).

[13] G. Birkhoff and Thomas E. Caywood, "Fluid flow patterns," NavOrd Report 447 (1947).

# Dust—An Active Meteorological Factor in the Atmosphere of Northern Africa

M. G. El-Fandy
*Assistant Professor, Farouk I University, Egypt*
(Received January 31, 1949)

In subtropical Africa dust is shown to be an active meteorological element, particularly below the *T*-inversions common at 2000–3000 feet under Etesian conditions. Actual instability of the lower atmosphere is then not necessary for dust to rise, and mere mechanical stirring by day by winds of about 20 miles per hour or more is sufficient to produce severe sand storms over desert areas under such conditions.

A dust particle suspended in the atmosphere has been assumed to possess an "air-pocket" or "surrounding eddy" of its own. The radius of the "air-pocket" of a particle of radius "$a$" cm (less than 5 micron) is shown to be about $10a$ cm. It is doubtful whether one can assume the persistence of one and the same "eddy," but with such an assumption made the average radius of the "eddy" is found to be $6a$ by observation, but on rejecting the assumption there appears to be a close agreement between theory and observation. The particles also share their excess of heat with their environment in a very small time, of the order of $10^{-2}$–$10^{-3}$ second.

The amount of heat that can be directly supplied to the lower atmosphere by baked dust blown in a dust storm may, by itself, be sufficient to produce convection up to a height of 3000 feet or more. The same range of height of convection is also attainable in a cloud of dust by the absorption of solar radiation scattered or diffusely reflected within the cloud. In this respect it is suggested, though not fully verified, that the effect of suspended dust should be taken into account when discussing problems of upper-layer instability in subtropical Africa.

It has also been observed that in the special cases of "squalls" the intensity of the storm, notably the poorness of visibility, is proportional to the time during which the squall reaches its maximum. The feature has been given a mathematical explanation.

The conclusion is that owing to increased absorption of solar radiation within dusty atmospheres, or to direct heating of the air by baked dust that may be blown by day in the warm season of the region, dust is found to play a part that cannot be ignored in modifying the processes of water vapor and heat transfer, by altering turbulent and radiative processes, and in modifying the properties of air masses invading the deserts.

## INTRODUCTION

PARTICLES of dust suspended in the atmosphere can have diameters larger than 0.1 micron but not less.[1] Samples of particles from layers of "dust-haze" or "drifting dust" at various heights, ranging from near the earth's surface up to about 3 km, were collected in Petri dishes and were examined under the microscope. The maximum diameters of the particles so collected could be taken as 4 micron.

Table I gives the limiting velocity of fall $V$ of dust particles relative to still air, using Cunningham's formula[2] and taking the density of dust to be 2.64. Although in a turbulent atmosphere conditions may be different from those assumed relative to still air, yet it is clear from the table that the particles liable to remain in suspension for a long time over a wide area must have radii less than 5 micron.

During the warm season of the region (March to September) the establishment of a *T*-inversion or reduced lapse-rate at about 2000 to 3000 feet above the earth's surface, together with the increase of surface wind above a critical value of about 20 miles per hour will cause the development of local sand storms which may extend vertically to 1000 feet or more.

TABLE I. Limiting velocity of fall of dust particles relative to still air.

| Radius "$a$" $\mu$ | Limiting velocity of fall $V$ cm/sec. | Limiting velocity of fall meters/day |
|---|---|---|
| 0.1 | 0.0007 | 0.6 |
| 0.5 | 0.0095 | 3.2 |
| 1.0 | 0.0346 | 30 |
| 2.5 | 0.2035 | 176 |
| 5.0 | 0.797 | 690 |

An example of this type occurred over the Western Desert on 30th April, 1941 and visibility was reduced to 100 yards. Figure 1 gives the dry-bulb readings from Sidi Barrani on that day. Surface heating by day together with heating of the surface air by baked blown dust soon established a super-adiabatic lapse-rate to about 2000 feet, as shown in Fig. 1, and the intensity of the storm was thus increased.

This type of sand storm is also much activated when local conditions are in favor of strong surface winds. An outstanding area is the Upper Nile Delta and Cairo District, where the "orographic bottle-neck effect" between the Mokattam Hills and the hills east and northeast of Faium (Fig. 2) favors great increase of wind speed near Cairo with S and SSW winds.

It will be seen, however, that owing to absorption of solar radiation within clouds of dust, or to direct heating of the lower atmosphere by baked dust that may rise by day, the eddy diffusivity $K$ is highly increased. Thus, representing $K$ by the usual notation[3]

$$K = \tfrac{1}{2}wd, \qquad (1)$$

[1] G. C. Simpson, Quart. J. Roy. Met. Soc. **67**, 115 (1941).
[2] E. Cunningham, Proc. Roy. Soc. London **83**, 357 (1910).
[3] D. Brunt, *Physical and Dynamical Meteorology* (Cambridge University Press, London, 1939), pp. 224–237.

where $w$ is the term for the vertical velocity and $d$ is the vertical extent since:

(1) In severe dust storms, such as "haboobs" or severe "khamsin"[4] dust may rise to about 3000 meters in an average of about 3 hours, $K$ amounts to about $10^6$ cm²/sec.

(2) In the common cases of air layers of "drifting dust" or "dust-haze" upper air ascents show that such atmospheres usually have sharp tops and that the top of a dusty layer usually undergoes an elevation by day. This elevation amounts to about 500 m or more in 6 hours during the warm season of the region, and is due to increased absorption of solar radiation within the dusty layer. The increase of turbulence and consequent redistribution of dust in the vertical by day also produce an apparent elevation of the base of the dusty layer, of the type shown in Table II, which gives the hourly surface observations of horizontal visibility at Almaza (Cairo District) on 3rd July, 1941. Dust was apparently elevated from the surface at about 0900 G.M.T., but came down again by 1630 G.M.T.

The upper air report from the same district (Heliopolis) on the same day stated: "Cloudless sky. Haze-top between 0300 and 0400 G.M.T. at 9500 feet. Horizontal visibility good at 10,000 feet." Between 0855 and 1025 G.M.T. another report stated: "Poor visibility at 10,000 feet. Haze-top at 11,000 feet." In other words, the haze-top was elevated by 450 m in 6 hours, so that $K$ amounts to $5\times10^4$ cm²/sec. in this particular case.

(3) With sand storms below $T$-inversions or reduced lapse-rates dust rises to about 2000 feet in about 3 hours, so that $K$ amounts to about $10^5$.

## ROLE OF DUST PARTICLES IN TURBULENT TRANSFER OF HEAT AND MOISTURE

A solid sphere moving through the atmosphere is assumed to be associated with a thin film of air. Pressure field is only disturbed within this film, and so long as the sphere is rising it must have an acceleration greater than the earth's acceleration $g$. If "$a$" is the radius of the solid sphere, the pressure "$p$" at any distance "$r$" from the center of the sphere is given by Stokes' approximation[5]

$$p=\text{constant}+\tfrac{3}{2}\mu U(a\cos\theta)/r^2, \qquad (2)$$

where $-U=$velocity of uniform stream of air of viscosity $\mu$,

$$=-2/9(\sigma-\rho)(ga^2/\mu),$$

in which $\sigma$ and $\rho$ are respectively the density of sphere and air.

Let $z=r\sin\theta$ and $x=r\cos\theta$, so that $p=$constant $+\tfrac{3}{2}\mu U(az/r^3)$, and

$$(\partial p/\partial z)=(3/2)\mu U(a/r^3)-(9/2)\mu U(az^2/r^5).$$

The boundary layer is defined such that at its outer surface $\partial p/\partial z$ must be equal to $-g\rho$, so that

$$(r/a)^3=(\sin^2\theta-\tfrac{1}{3})(\sigma-\rho)/\rho, \qquad (3)$$

which represents two double solids (Fig. 3), the total volume of which is $Q$ and is equal to the volume of the boundary layer.

Therefore,

$$Q=2\int_0^s \pi z\,dx+2\int_0^{2s}\pi x^2\,dz,$$

where $s^3=\tfrac{1}{3}a^3(\sigma-\rho)/\rho$.

If "$b$" is the radius of a sphere whose volume is equal to $Q$, then $b$ equals about $13a$, and for all practical purposes

$$b=10a, \qquad (4)$$

is a fair approximation, taking $\sigma/\rho$ to be about 2000 for dust in air.

It will now be assumed that every boundary layer is a physical representation of an "eddy." In other words, it will be assumed that one and the same physical process is responsible for turbulent transfer of water vapor and dust particles. Usually layers of air of relatively high moisture and dust content exist below temperature reduced lapse-rates or inversions, such as those formed by subsidence within the northerly "Etesian" winds of summer.

Let $n$ be the average number of dust particles per cc of air (=concentration) at a certain level which has become rich both in water vapor and dust content. The mass of water vapor transferred

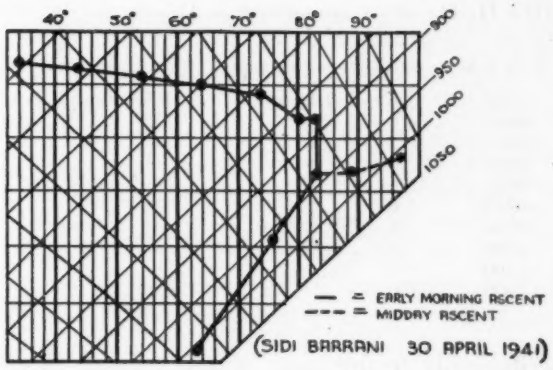


FIG. 1. Sidi-Barrani dry-bulb readings on tephigram (Western Desert 30/4/1941).

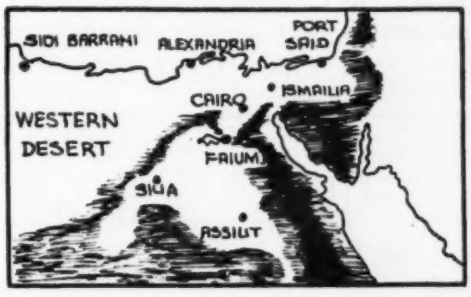


FIG. 2. "Orographic bottle-neck effect" between Mokattam Hills and the hills E and NE of Faium.

[4] M. G. El Fandy, Quart. J. Roy. Met. Soc. 66, 323–336 (1940).

[5] L. Hoph, *Handbuch der Physik* (Verlag Julius Springer, Berlin, 1927), Bd. 7, 2, pp. 107–108.

TABLE II. The dusty atmosphere of Almaza, 3rd July, 1941.

| Time G.M.T. | Visibility in code figure | Present weather in code |
|---|---|---|
| 0345 | 6 = (4–10 km) | 05 = (dust haze) |
| 0445 | 6 = | 05 |
| 0545 | 6 | 05 |
| 0645 | 6 | 05 |
| 0745 | 6 | 05 |
| 0845 | 6 | 05 |
| 0945 | 8 = (20–50 km) | 00 = (clear) |
| 1200 | 8 | 00 |
| 1500 | 8 | 00 |
| 1800 | 6 | 05 |

mechanically to the layer by the vertical diffusion of dust will be $4/3\pi b^3\rho_1 n x_1$ gram per unit volume, where $\rho_1$ is the air density and $x_1$ the average value of the humidity mixing ratio (in grams per gram of air) near the earth's surface during the rise of dust. The net increase of water vapor within the level will be $4/3\pi n b^3(x_1\rho_1 - x_2\rho_1)$ gram per unit volume, where $x_2$ is the original average value of $x$ within that level (i.e., before the addition of dust). Let the increase per gram of air be denoted by $\Delta x$, so that

$$\Delta x = 4/3\pi n b^3(x_1 - x_2)\rho_1/\rho_2. \qquad (5)$$

To compare this theory with observation, consider the dusty atmosphere of 3rd July, 1941, which prevailed over Egypt after sand storms of the above described reduced lapse-rate type over the Western Desert on July 2, 1941. Heliopolis' early morning tephigrams of the 2nd and 3rd together with the mid-day dry-bulb readings of the 2nd are reproduced in Fig. 4. From these tephigrams the curves of Fig. 5 have been deduced to find the vertical distribution of $x$ (per kilogram of air in this case) and the wet-bulb potential temperature $\theta'$. Continuous curves such as $ABCD$ represent the early morning distribution of the 2nd, and dotted lines, such as $A'B'C'D'$, represent the corresponding distribution of the 3rd. These curves show clearly that although the same air mass prevails between 2000 and 10,000 feet on both days, yet there is an average increase in $x$ by 0.8 gram and an average increase in $\theta'$ by about 3.5°C within 24 hours.

The turbulent transfer of water vapor by air and by particles of dust takes place in proportion to their masses, so that on allowing for proper units and applying Eq. (5) to the above example we have:

$$\Delta x = 0.8(4/3\pi a^3 n)\sigma/\rho,$$

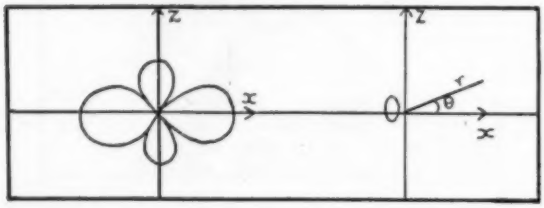


FIG. 3. The two double solids of Eq. (3).

where $\sigma/\rho = 3000$ in this case, $x_1 = 13.5$, $x_2 = 4.5$, $\rho_1/\rho_2 = 1.08$. Therefore, $b^3 = 240a^3$ or $b = 6a$ approximately.

On the other hand, if instead of giving $x_1$ the value near the earth's surface only it is given an average of 7, which in our example corresponds to the whole path in the turbulent atmosphere, we get for a fair approximation

$$b = 10a, \qquad (6)$$

as before, which shows a close agreement between theory and observation.

Again, when a dust particle at any temperature is blown into the air a rapid process of sharing heat between the particle and its environment takes place. The thermal capacity $H_1$ of a dust particle of radius "$a$" and specific heat $c$ together with its accompanying pocket of average radius "$10a$" is

$$H_1 = 4/3\pi a^3(\sigma c + Cp),$$

where $Cp$ = sp. ht of air at constant pressure = 0.24 cal./g.

$$1000\rho = 1 \text{ approximately,}$$

and

$$c = 0.188 \text{ cal. per gram.}$$

As an example let "$a$" be given the convenient value $10^{-4}$ cm, so that

$$H_1 = \pi \cdot 10^{-12} \text{ unit.}$$

Let the particle be originally in an environment of temperature $T_d$ and be suddenly projected into another environment of temperature $T_a$, where $T_d$ is greater than $T_a$. (This is almost always the case when dust is blown by day.) In a small interval of time $dt$, the amount of heat conducted away from the particle to its new environment will be $H_2$ where, for a first approximation

$$H_2 = 4\pi a^2 k(q/a)dt \text{ unit,}$$

in which $k$ is the air-conductivity ($= k'\rho Cp$; $k'$ being

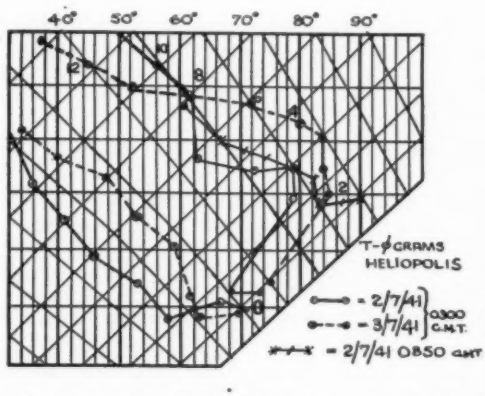


FIG. 4. Early morning Heliopolis tephigrams of 2nd and 3rd July, 1941, and mid-day dry-bulb readings of 2nd.

the molecular conductivity), which can be taken as $6\times10^{-5}$ c.g.s. unit, and $q$ is the difference between the temperature of the particle and that of the outer atmosphere during the time interval $dt$.

The average temperature difference between the particle and its new environment during the time $t$ required for thermal equilibrium is $(T_d-T_a)/2$, so that in our example above the total heat that can be transferred by conduction is

$$\int_0^t H_2dt=\pi(T_d-T_a)t\cdot10^{-8}$$

unit approximately.

The amount of heat carried out from the boundary layer by the process of replacement during the time $dt$ is

$$H_3=4\pi b^2F\rho Cpq'dt,$$

where $q'$ is the number of °C of the boundary layer above the new environment during the interval $dt$, and $F$ is the rate at which air within the boundary layer may be replaced by normal air from the turbulent region outside (in cc per unit area).

After time $t$ the particle and its pocket become in thermal equilibrium with their new environment, and the total heat conducted by replacement will be

$$\int_0^t H_3dt=1/2(T_d-T_a)Ft\cdot10^{-9}$$

approximately, where $q'$ is given the same average value as $q$.

Equating heat lost and heat gained during such processes we get

$$H_1(T_d-T_a')=\int_0^t(H_2+H_3)dt \qquad (7)$$

where $T_a'$ is the final temperature attained by the dust particles and environment.

In the example under consideration we have

$$(T_d-T_a')\cdot10^{-12}=(T_d-T_a)t\cdot10^{-8}+\tfrac{1}{2}(T_d-T_a)Ft\cdot10^{-9},$$

and on collecting terms in $t$ we get, for a first approximation,

$$t=(T_d-T_a')[(T_d-T_a)(10+\cdot5F)]^{-1}\cdot10^{-3} \qquad (8)$$

which shows that for measurable temperatures in the atmosphere and for all values of $F$, $t$ is only a very small fraction of a second.

Aitken[6] also made the observation that dust motes are not affected by solar heat focused by a large lens, while larger objects burn up at once under the same conditions. The explanation was that the mote lost heat very rapidly to the environment.

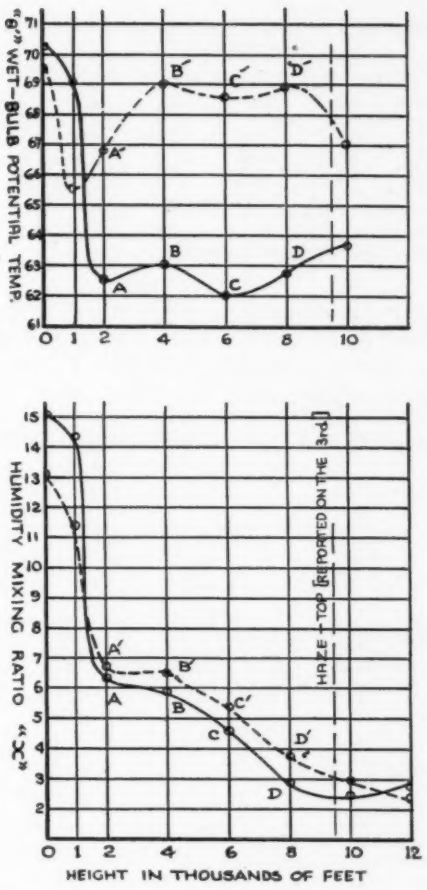


FIG. 5. Vertical distribution of $x$ and $\theta'$ from Heliopolis early morning ascents on 2nd and 3rd July, 1941. $ABCD$= curves for 2nd. $A'B'C'D'$=curves for 3rd.

## ROLE OF DUST PARTICLES IN DEVELOPMENT OF CONVECTION

The relation between the concentration $n$ at any level $z$ and the limiting velocity of fall $V$ to the degree of turbulence in the air is given by[7]

$$\frac{\partial n}{\partial t}=\frac{\partial}{\partial z}\left(K\frac{\partial n}{\partial z}\right)+V\frac{\partial n}{\partial z}. \qquad (8)$$

Upper air observations show that in dusty atmospheres practically dry-adiabatic lapse-rates are established, except for reduced lapse-rates or isothermal distributions within the first 1000 to 2000 feet, so that $K$ can be treated as a constant independent of height within dusty atmospheres and

$$\partial n/\partial t=K(\partial^2n/\partial z^2)+V(\partial n/\partial z). \qquad (9)$$

[6] J. Aitken, Trans. Roy. Soc. (Edinburgh) **42**, 489 (1902).

[7] C. W. Normand, Indian Met. Office (Calcutta) 22, VI (1921).

TABLE III. Lower Egypt monthly actinometric averages in °C.*

| | March | April | May | June | July | Aug. | Sept. |
|---|---|---|---|---|---|---|---|
| (a) | 62.3 | 67.2 | 70.0 | 72.9 | 72.4 | 71.9 | 69.4 |
| (b) | 41.3 | 46.2 | 49.6 | 52.1 | 52.0 | 51.3 | 49.3 |
| (c) | 31.0 | 36.0 | 39.6 | 42.2 | 41.9 | 41.4 | 39.6 |

* (a) = Black-bulb in vacuum; (b) = Bright-bulb in vacuum; (c) = Maximum air-temperature in standard screen.

Under a steady state (when $\partial n/\partial t=0$), let the top of the "haze-layer" be at a height $h$ above an arbitrary chosen base and put $V/K=B$. The boundary conditions are

$n=n_0$ at $z=0$ and $n=\alpha n_0$ at $z=h$, where $0<\alpha\leqslant 1$.

Therefore,

$$n=\frac{n_0[\alpha-e^{-Bh}+(1-\alpha)e^{-Bz}]}{1-e^{-Bh}}. \tag{10}$$

Integrating from $z=0$ to $z=h$ in order to get the total number $N$ of dust particles contained within a vertical cylinder of unit cross section we have

$$N=\int_0^h n\,dz$$

$$=\frac{n_0[(\alpha-e^{-Bh})h+(1-\alpha)(1-e^{-Bh})/B]}{1-e^{-Bh}}. \tag{11}$$

When, for example, the range of horizontal visibility at the base is half that at the top, $\alpha=\frac{1}{2}$. Usually $K$ is large and in dust storms it approaches the order of $10^6$ or $10^7$ indicating active turbulence and $\alpha=1$, so that $N=n_0h$.

The thermal capacity $C$ of $N$ particles of dust is $4/3\pi a^3\sigma CN$, and on making the following substitutions for a moderate dust storm: $h=3\times10^5$ cm, $a=2$ micron, and $n=10^6$, we get $C=5$ units.

It is evident from Table III that the maximum temperatures attainable by desert sand during the warm season of the region are on the average 25°C above those attainable by the surface air, so that the maximum amount of heat that may be supplied to the lower atmosphere by baked blown dust is 125 cal. per cm².

Adopting the principle used by Gold[8] in calculating the height of convection resulting from the addition of the above heat energy to the lower atmosphere, we see that the height of such convection amounts to 3300 feet. In other words, the energy so supplied by dust is sufficient to create convection up to more than one-third of the assumed height of the storm. This may be taken as an explanation of the commonly observed feature that the warm season dust storms of the region have a marked tendency to attain their maxima in the early afternoon.[9]

In Egypt accurate readings of desert temperature are not available, but regarding dust as intermediate between black and bright, such readings can be taken to lie between the values given in the rows marked (a) and (b) in Table III, in which monthly maximum actinometric averages in °C for lower Egypt are reproduced.[10] The readings marked (a) are of black-bulb in vacuum, while (b) are of bright-bulb in vacuum. Maximum air temperatures in standard screen are given in (c).

Returning now to the dusty atmosphere of 3rd July, 1941, in which the top was elevated by 1500 feet through absorption of solar energy, we see that if such suspended dust is considered to be in vacuum (i.e., isolated from the air to prevent conduction of heat) its temperature will on the average be not less than 25°C higher than the temperature of the air in which it is suspended. In other words, $C=1$ unit in this case and corresponds to a concentration $n=2\times10^5$ particles of radius 2 micron per cc of air.

Since upper layer instability may pass into the thunder-storm stage by convection during the day, the presence of sufficient dust in the atmosphere of subtropical Africa should introduce some modification to our present conception regarding this problem of instability. However, it was suggested by Farquharson[11] that the brownish shade nearly always observed at the base of the high cumulus cloud at Khartoum is evidence that even on days of light winds upward drift of sand occurs by day.

## ROLE OF DUST IN RADIATION BALANCE

As half of the total energy in the radiation from the sun can be taken as contained within the range $0.4\mu$ to $0.7\mu$, i.e., in the form of light, the average diameter of suspended particles is large compared with these radiations. Diffuse reflection takes place and Rayleigh's argument[12] can be used.

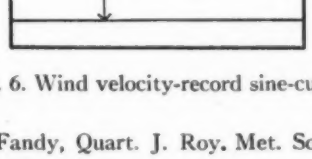


FIG. 6. Wind velocity-record sine-curve.

[8] E. Gold, Professional Notes No. 63, 5–6 (1933).

[9] M. G. El-Fandy, Quart. J. Roy. Met. Soc. 66, 323–336 (1940).

[10] Ministry of Public Works, 1938, Climatological Normals for Egypt and the Sudan, Cairo, Egypt.

[11] J. S. Farquharson, Quart. J. Roy. Met. Soc. 63, No. 271 (1937).

[12] W. J. Humphreys, *Physics of the Air* (McGraw-Hill Book Company, Inc., New York, 1929), pp. 573–577.

On the other hand, the point of maximum intensity in the terrestrial radiation, with an average temperature of 30°C, has a wave-length of about 10$\mu$. In other words, the greater part of the effective wave-length of the outgoing radiation is greater than 5$\mu$. Suspended dust, therefore, mainly scatter radiations from the earth. However, it is made clear by Humphreys that a veil of dust is more effective in shutting solar radiation out than it is in keeping terrestrial radiation in.

Assuming the dust particles to be full radiators at their temperatures, consider a unit volume of unit cross-sectional area and let the maximum effective vertical radiation from the contained dust be

$$2\pi a^2 n(ST'^4),$$

where $S$ is Stefans' constant. Let $T''$ be the effective temperature at which the particles are absorbing heat. The energy lost per unit volume will be

$$2\pi a^2 nS(T'^4 - T''^4),$$

and the rate of cooling $\Delta T/\Delta t$ in °C per second will be given by

$$(\Delta T/\Delta t)\rho Cp = 2\pi a^2 nS(T'^4 - T''^4). \qquad (12)$$

For all reasonable values of the terms in Eq. (12) this rate of cooling is negligible, and in northern Africa suspended dust is only a factor that increases the heat energy of the lower atmosphere.

Bearing such factors in mind, consider a cloud of dust within which 80 percent of the solar radiation is ultimately absorbed and 10 percent of the outgoing terrestrial radiation is also absorbed. The amount of solar radiation available in the lower atmosphere can be taken as 60 percent of the total amount received in the outer atmosphere as calculated by Angot.[13] Column 1 of Table IV gives the appropriate amount so received at latitude 30°N (in cm² on form 2810 tephigram, Air Ministry Meteorological Office, London). 80 percent of this amount is ultimately absorbed within the cloud and is represented in column 2. Columns 3 and 4 give respectively the average effective temperatures at which the earth's surface radiates and 10 percent of this energy in cm² on same tephigram. Column 6 represents half the amount of heat radiated to space by the cloud of dust, assuming the particles to be 70 percent full radiators at the effective temperatures given in column 5. The other half, or even much more, may be taken to be absorbed. In column 7 is given the net energy gained by the cloud in cm² on the tephigram per day.

## MODIFICATION OF AIR MASSES BY DESERT TRAJECTORY

An almost continuous process of adding dust and sand to the atmosphere of northern Africa takes place through dust storms or dust-raising winds and causes heating of the lower atmosphere. The net gain in energy given in column 7 of Table IV corresponds to an increase in temperature of the magnitude indicated in column 8.

Heating is a maximum in the summer months and is almost nil in December. The practical significance of such results lies in estimating how far air currents invading subtropical Africa and causing dusty atmospheres of any form become modified. The feature usually takes place on a large scale, and in the warm season air currents are highly modified by desert trajectory. For example, polar air that crosses Egypt takes about 4 to 5 days to reach the central Sudan and its temperature is then not less than 10°C higher than over Egypt.

TABLE IV. Heating of air masses by clouds of dust.

| Month | 1[a] | 2[b] | 3[c] | 4[d] | 5[e] | 6[f] | 7[g] | 8[h] |
|---|---|---|---|---|---|---|---|---|
| Jan. | 13.3 | 9.6 | 290 | 3.7 | 280 | 11.5 | 2 | 0.7 |
| Feb. | 16.8 | 12.1 | 295 | 4.0 | 285 | 12.2 | 4 | 1.3 |
| March | 20.3 | 14.6 | 300 | 4.3 | 290 | 13.0 | 6 | 2.0 |
| Apr. | 23.5 | 16.9 | 305 | 4.6 | 295 | 14.0 | 7.5 | 2.5 |
| May | 24.7 | 17.8 | 310 | 4.9 | 300 | 15.0 | 8 | 2.7 |
| June | 26.7 | 19.2 | 315 | 5.2 | 300 | 15.0 | 9 | 3.0 |
| July | 25.3 | 18.2 | 320 | 5.5 | 300 | 15.0 | 9 | 3.0 |
| August | 23.5 | 16.9 | 315 | 5.2 | 300 | 15.0 | 7 | 2.3 |
| Sept. | 21.3 | 15.4 | 310 | 4.9 | 300 | 15.0 | 5 | 1.7 |
| Oct. | 17.3 | 12.5 | 300 | 4.3 | 295 | 14.0 | 3 | 1.0 |
| Nov. | 14.2 | 10.2 | 295 | 4.0 | 290 | 13.0 | 1 | 0.3 |
| Dec. | 12.3 | 8.9 | 290 | 3.7 | 285 | 12.2 | 0 | 0.0 |

[a] 60 percent of the solar radiation at lat. 30°N in cm² on Form 2810, (Air Ministry of London) tephigram.
[b] 80 percent of column 1.
[c] Effective temperature of earth's radiation in °A.
[d] 10 percent of radiation at temperatures given in column 3, represented cm² on tephigram.
[e] Effective temperature of dust in °A.
[f] ½ radiation from dust in cm² on tephigram.
[g] Net energy gained by cloud of dust.
[h] Rise of temperature in °C per day.

## APPENDIX

## Some Characteristics of Surface Wind in Dust Storms

The intensity of a dust storm, in particular the poorness of visibility, is observed to be governed not only by the increase of horizontal wind speed above a certain critical value, but also by the time interval that is seen to lapse in attaining a velocity maximum. The longer the wave on the wind velocity record the more intense the storm is. In a single marked squall that continues for an appreciable time, say, 15 minutes (which is the usual case with instability conditions), visibility falls rapidly while extensive rise of dust occurs.

It is natural to infer to the increase in the vertical component of wind the variation in the intensity of the storm, or the variation in the density and vertical extent of suspended dust. Let $\tau$-seconds denote the period of a single wave on the wind record and assume a sine-curve for the velocity record as represented in Fig. 6. Also let the time be measured from $O$, so that time at $O$ is zero.

At any time $t$ let the horizontal wind velocity be represented by $(u+u')$, where $u'$ is the departure from the average value $u$. Denote the average maximum departure from $u$ by $(u')_M$ as represented in Fig. 6, and let $(w, w', (w')_M)$ represent the corresponding values for the vertical component. The increase

[13] See reference 3, p. 112.

in the kinetic energy of a particle of mass $m$ from time $t$ to time $t+dt$ is

$$\tfrac{1}{2}md\{(u+u')^2+(w+w')^2\},$$

and this is equivalent to a fraction $1/f$ of the energy taken in lifting the particle through a height $(w+w')dt$, where $f>1$, since some of the energy for lifting is taken from the air, so that

$$mg(w+w')dt=\tfrac{1}{2}mfd\{(u+u')^2+(w+w')^2\}$$

or

$$(w+w')=(f/2g)(d/dt)\{(u+u')^2+(w+w')^2\}.$$

The average value of $(w+w')$ within the interval $-\tau/4$, $+\tau/4$ is such that

$$w=\frac{1}{\tau/4}\int_{-\tau/4}^{+\tau/4}(w+w')dt=\frac{f}{\tau g}\int_{-\tau/4}^{+\tau/4}d\{(u+u')^2+(w+w')^2\}$$
$$=4f/g\tau[u(u')_{Av}+w(w')_{Av}].$$

Therefore,

$$(w')_{Av}=(g\tau/4f)-uu'/w$$

and to a first approximation the maximum departure from the usual vertical component is given by

$$(w')_{Av}=g\tau/4f. \tag{13}$$

Equation (13) shows that the average maximum departure from the usual vertical component must increase with the period, and explains the observed feature.

## Some Factors Affecting the Rate of Precipitation Hardening in Cu-Be Alloys

PAUL A. BECK
*Department of Metallurgy, University of Notre Dame, Notre Dame, Indiana*
(Received December 22, 1948)

It was found that the lower peak hardness and faster overaging of a Co-containing Cu-Be alloy quenched from a low solution treating temperature is associated with a large amount of grain boundary precipitation. A similarly large amount of grain boundary precipitation, coupled with low peak hardness value, is obtained if the alloy contains a small amount of Cr impurity, even if the quenching temperature is normal. The grain boundary reaction resulting from the Cr impurity can be eliminated by using high solution treating temperatures.

IT has been known for a long time that the precipitation hardening process in Cu+2 percent Be is considerably affected by the addition of a few tenths of a percent of Co, Ni, or Fe. An outstanding effect of these additions is the considerable decrease in the rate of hardening and particularly in the rate of overaging at a certain temperature. This is shown in Fig. 1, where curve 1 represents the variation of hardness with aging time at 316°C for the binary Cu+2 percent Be alloy, and curve 2 gives similar data for an alloy with the same Be content, but containing in addition 0.38 percent Co. It is seen that, in comparison with the Co containing alloy, the binary alloy begins to harden at a considerably higher rate, reaches a lower peak hardness, and then rapidly overages. Comparison of the microstructures shows that in the binary alloy overaging is associated with the development and progress of the grain boundary reaction as described in a recent paper by Guy, Barrett, and Mehl.[1] As shown by these authors,

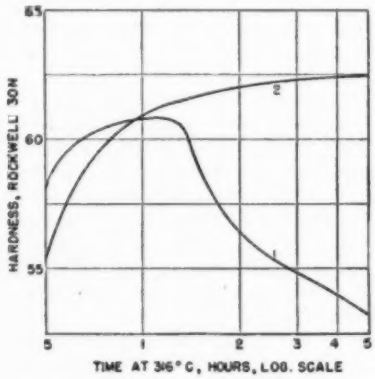


FIG. 1. Hardness *vs.* time at aging temperature (log scale) No. 2 hard strip. (1) Cu+2 percent Be alloy. (2) Cu+2 percent Be+0.38 percent Co alloy.

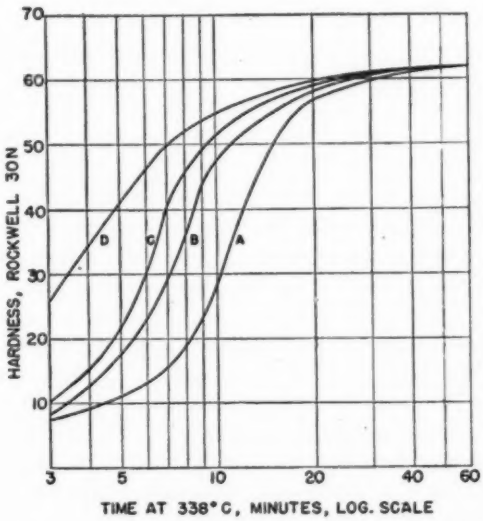


FIG. 2. Hardness *vs.* time at aging temperature (log scale) for Cu+2 percent Be+0.4 percent Co alloy after quenching from various temperatures (*A*) quenching temperature 760°C, (*B*) quenching temperature 787°C, (*C*) quenching temperature 815°C, (*D*) quenching temperature 842°C.

[1] Guy, Barrett, and Mehl, *Mechanism of Precipitation in Alloys of Beryllium in Copper*, AIME RP. 2341 (Metals Technology, February 1948).

the area affected by the grain boundary reaction has lamellar structure and lower hardness than areas with general precipitation. During overaging this area gradually extends to cover practically the whole specimen. Under otherwise similar conditions the extent of the grain boundary reaction in the Co containing alloy is very much lower. The structure appears to be spheroidal rather than lamellar. Although even here the affected area increases with time to some extent, normally it remains confined to the vicinity of the grain boundaries. This may explain the great stability of the Co containing alloy against overaging. The Co additions decrease the rate of both the general precipitation reaction and the grain boundary reaction.

It was noted that by varying the quenching temperature of a Co containing Cu-Be alloy, the rate and the mechanism of the precipitation reaction during subsequent aging can be changed considerably. Figure 2 gives hardness *vs.* aging time curves for four different quenching temperatures. Curve *A* represents hardening at 338°C after quenching from 760°C, which is considered the normal quenching temperature for this alloy. Curves *B*, *C*, and *D* indicate the effect of increasing quenching temperatures. It is seen that the maximum hardness is practically unaffected, but that the initial rate of hardening greatly increases with increasing quenching temperature. It is interesting to note that with increasing quenching temperature, the extent of the grain boundary reaction is greatly reduced. With quenching temperatures above about 787°C it practically disappears. That the suppression of the grain boundary reaction here is not a result of an increase in the amount of Co entering into solid solution is indicated by the fact that the initial rate of hardening is at the same time increased. This is just the opposite of what may be expected from an increase in the Co content (see Fig. 1). If the quenching temperature is decreased below 760°C, the result is a great increase in

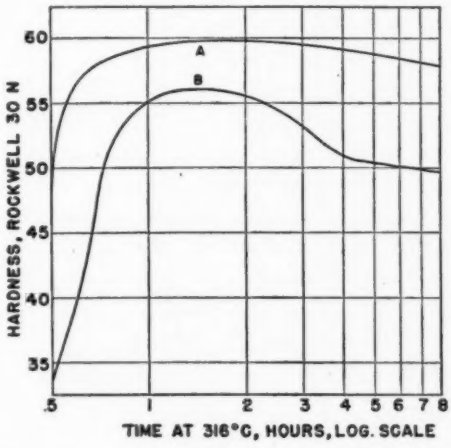


FIG. 3. Hardness *vs.* time at aging temperature (log scale) for Cu+2 percent Be+0.25 percent Co alloy quenched from (*A*) 760°C, (*B*) 726°C.

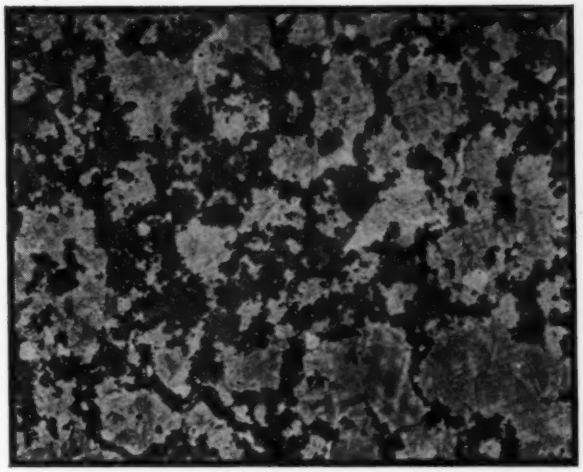

FIG. 4. Cu+2 percent Be+0.25 percent Co alloy containing Cr impurity. Quenched from 760°C, aged for 4 hours at 316°C ×150. Large extent of grain boundary reaction.

the amount of grain boundary reaction and, correspondingly, an increase in the rate of overaging. This is shown in Fig. 3, which indicates in addition that the initial rate of hardening, as well as the maximum hardness reached, is lowered by decreasing the quenching temperature below 760°C. Low quenching temperatures affect the rate of the grain boundary reaction and the extent of overaging in the same manner as a decrease in the Co content would, However, the initial rate of hardening is again affected in the opposite sense. Apparently, then, variations in the quenching temperature do not affect the rate and the extent of the grain boundary reaction by changing the effective amount of Co in solid solution.

For a given quenching temperature the rate of the grain boundary reaction in Co containing Cu-Be alloys was found to be greatly affected by the addition of Cr

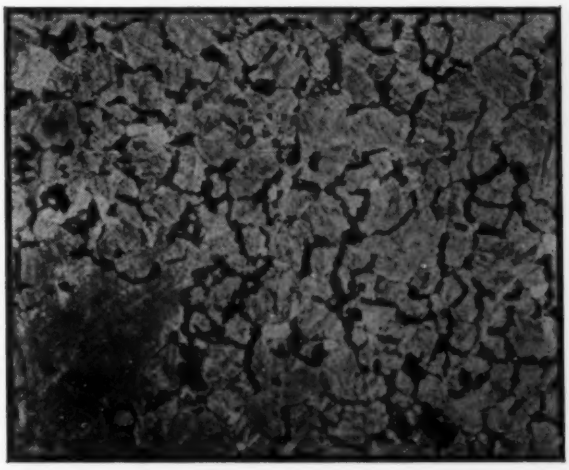

FIG. 5. Cu+2 percent Be+0.25 percent Co alloy containing Cr impurity. Quenched from 787°C, aged for 4 hours at 316°C ×150. Extent of grain boundary reaction decreased considerably.

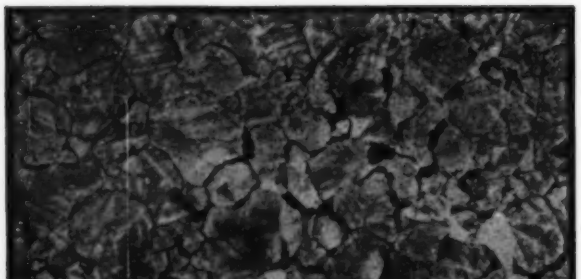

FIG. 6. Cu+2 percent Be+0.25 percent Co alloy containing Cr impurity. Quenched from 815°C, aged for 4 hours at 316°C×150. Extent of grain boundary reaction corresponds to that normally found in Cr-free alloy after quenching from 760°C and aging.

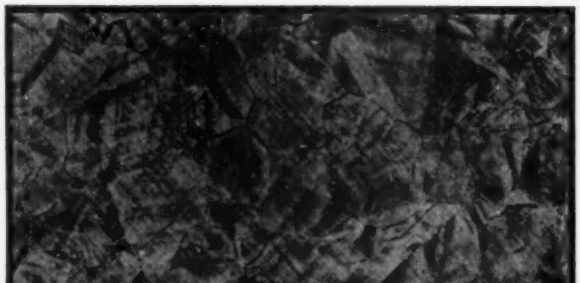

FIG. 7. Cu+2 percent Be+0.25 percent Co alloy containing Cr impurity. Quenched from 842°C, aged for 4 hours at 316°C×150. Grain boundary reaction is almost completely suppressed.

in spectrographic quantities. In an alloy quenched from 760°C, very small amounts of Cr will increase the rate of the grain boundary reaction to such an extent that the attainable maximum hardness no longer meets the standard specifications. However, it was found that in such Cr and Co containing alloys the grain boundary reaction may be progressively suppressed by using increasing quenching temperatures. This effect is similar to that described above for the Cr-free alloy. But the small Cr addition increases the quenching temperature necessary for suppressing the grain boundary reaction by as much as 50°C. Figures 4, 5, 6, and 7 show the decrease in the amount of the grain boundary reaction with increasing quenching temperature in an alloy containing 0.3 percent Co and a spectrographic quantity of Cr.

# Semi-Conducting Properties in Oxide Cathodes*

N. B. Hannay, D. MacNair, and Addison H. White
*Bell Telephone Laboratories, New York, New York*
(Received September 16, 1948)

It has been widely assumed, without adequate experimental verification, that barium-strontium oxide, as used in the oxide cathode, is an excess electronic semi-conductor. Accordingly, the electrical conductivity of (Ba, Sr)O has been studied as a function of temperature before and after activation with methane, extensive precautions being taken to exclude spurious effects. The increase in conductivity obtained characterizes (Ba, Sr)O as a "reduction" semi-conductor, and hence very probably as an electronic semi-conductor whose conduction electrons arise from a stoichiometric excess of (Ba, Sr) atoms in solid solution.

A basic prediction of the semi-conductor theory has been tested quantitatively with the finding that the electrical conductivity and the thermionic emission of a (Ba, Sr)O cathode are directly proportional through three orders of magnitude of activation; well-defined chemical and electrical activation and deactivation procedures were used in obtaining this result. It may be concluded that activation represents an increase in the chemical potential of the electrons in the oxide, little or no change in the state of the surface occurring. It has also been found that deviations from the proportionality of conductivity and emission may be expected under conditions leading to inhomogeneity in the oxide, in agreement with the semi-conductor theory also.

## I. INTRODUCTION

THE massive literature of the oxide-coated cathode shows it to be a very complex system which is by no means completely understood.[1] It is clear that the most nearly satisfactory theory of the emission of this cathode is based on the assumption that barium-strontium oxide (hereafter designated as (Ba, Sr)O) is an electronic semi-conductor activated by an excess of, e.g., barium dissolved in it. Experiment is generally consistent with this semi-conductor picture, under conditions where the thermal equilibrium of electrons assumed for the theory is likely to be realized.[1,2]

The semi-conductor theory of thermionic emission predicts a correlation between the thermionic emission and the electrical conductivity if the oxide is an homogeneous excess electronic semi-conductor. If there is no change in the state of the surface during activation, the conductivity and emission should be directly proportional. A primary object of the present investigation has been to test this basic prediction by studying the emission and conductivity as a function of the state of activation of the oxide, and thus obtain more information about the activation process than has hitherto been available. The theory underlying this work will be discussed briefly (Section II).

The basic assumption that (Ba, Sr)O is an excess electronic semi-conductor has not, however, been adequately tested by experiments specifically designed for the purpose. Accordingly, the initial phase of the investigation has been to test this assumption by studying the effect of high temperature treatment with a chemical reducing agent, methane, on the electrical conductivity of the oxide. A marked increase of conductivity under such treatment will characterize (Ba, Sr)O as a "reduction" semi-conductor, and thus very probably as an excess electronic semi-conductor.[3,4]

## II. SEMI-CONDUCTOR MODEL FOR OXIDE CATHODE

The theory is available in the literature,[1,5,6] and so only the part pertinent to this work will be briefly reviewed.[7]

Without making any explicit assumptions about the exact model to be used, it may be shown easily that the conductivity and emission should be related. The conductivity will be approximately proportional to the concentration of "free" electrons in the conduction band of the solid, $n_f$, at any fixed temperature.

$$\sigma = k_1 n_f. \tag{1}$$

The equilibrium thermionic emission current will be proportional to the concentration of conduction electrons and to $e^{-(\chi/kT)}$, $\chi$ being the work function for these electrons; the conductivity-emission relationship then follows immediately.

$$i = k_2 n_f e^{-(\chi/kT)}, \tag{2}$$

$$= k_3 \sigma e^{-(\chi/kT)}. \tag{3}$$

* A preliminary report of this work was made to the American Physical Society, Division of Electron and Ion Optics; Phys. Rev. **72**, 153 (1947).

[1] J. P. Blewett, J. App. Phys. **10**, 668, 831 (1939); *ibid.*, **17**, 643 (1946).

[2] J. A. Becker and R. W. Sears, Phys. Rev. **38**, 2193 (1931); W. Albricht, Physica **11**, 146 (1931); A. L. Reimann and L. R. G. Treloar, Phil. Mag. **12**, 1073 (1931); W. Meyer and A. Schmidt, Zeits. f. tech. Physik **13**, 137 (1932); E. Nishibori and H. Kawamura, Proc. Phys. Math. Soc. Japan **22**, 378 (1940); A. Fineman and A. Eisenstein, J. App. Phys. **17**, 663 (1946).

[3] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Clarendon Press, Oxford, 1940), Chapter 5.

[4] Methane has been reported to activate the electron emission of (Ba, Sr)O which contains finely divided nickel. (See C. H. Prescott and J. Morrison, J. Am. Chem. Soc. **60**, 3047, 1938). These authors do not discuss the role of conductivity.

[5] R. H. Fowler, *Statistical Mechanics* (University Press, Cambridge, 1936), second edition, pp. 401, 419; see reference 1, (1939).

[6] Frederick Seitz, *Modern Theory of Solids* (McGraw-Hill Book Company, Inc., 1940).

[7] The authors are indebted to Dr. C. Herring for many helpful discussions of the theory.

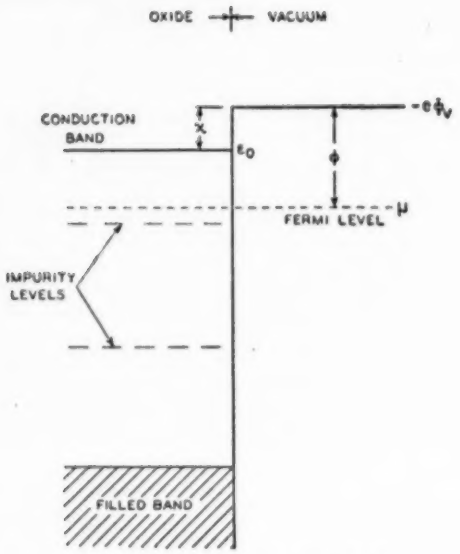


FIG. 1. Energy levels in (Ba, Sr)O.

The direct proportionality between the conductivity and emission expressed by Eq. (3) can be derived also in a more explicit form by the usual type of mathematical treatment, while maintaining the generality of the argument. In Fig. 1 we let the energy of an electron just outside the semi-conductor be $-e\Phi_v$, and at the bottom of the conduction band $\epsilon_0$, the zero of energy being taken at the average electrostatic potential inside the crystal, and $\Phi_v$ being the electrostatic potential in the vacuum just outside the semi-conductor. The quantity $\chi$ is the work function for conduction electrons used above, or the "electron affinity" of the crystal, and the quantity $\mu$ is the chemical potential of the electrons[8] (the partial molar-free energy of the electrons), or the "Fermi level." Impurity levels are shown, which may be of the *F*-center type, but their nature need not be specified for this discussion. The conductivity is then given by the expression

$$\sigma=(8e^2l_0/3h^3)(2\pi m^*kT)\exp[(\mu-\epsilon_0)/kT], \quad (4)$$

where $l_0$ = the electronic mean free path, $m^*$ = effective mass of the electron, and the other symbols have their usual meanings. The thermionic emission current under equilibrium conditions is given by an expression which may be derived either from thermodynamics or from statistical mechanics, and which may be written as

$$j=[(2\pi m^*kT)2kTe/h^3](1-\bar{r})\times\exp[(\mu-\epsilon_0-\chi)/kT], \quad (5)$$

where $\bar{r}$ = a reflection coefficient, and the true thermionic work function is given by

$$\varphi=-[\mu-\epsilon_0-\chi]. \quad (6)$$

The chemical potential, of course, will have definite meaning regardless of the model chosen. For a particular model it may be written explicitly in terms of the impurity centers, but this is not necessary for purposes of the present discussion. It is shown in Fig. 1 in the vicinity of the uppermost impurity levels, but could be drawn elsewhere without affecting the validity of the argument.

Upon activation of the semi-conductor, the chemical potential will increase, for example, by an increase in the number of impurity levels, and the increase in the chemical potential will increase both the conductivity and the thermionic emission, as shown by Eqs. (4) and (5). During this activation, if no quantities other than $\mu$ on the right hand sides of (4) and (5) change, the conductivity and emission will be directly proportional. If, on the other hand, $\chi$ for example should change because of a change in the surface dipole, then this linear relationship would no longer hold. It should be emphasized that the validity of these predictions does not depend on a particular model for the semi-conductor.

Results of Heinze and Wagener in electron microscope studies tend to indicate that the chemical potential does change during the activation process, with little change in the surface dipole.[9,10] Nishibori and Kawamura, Albricht, and others[2] have measured the conductivity and emission of an oxide cathode, with the work of the first of these as perhaps the best experimental check, a linear relationship being derivable from their three experimental points. The general inadequacy of the data made it appear desirable to investigate the conductivity-emission relationship much more exhaustively, and under activation procedures in which chemical and electrical processes were more clearly defined.

## III. CONDUCTIVITY MEASUREMENTS

A condensed description of the experiments performed and the results obtained will first be presented, and will be followed in Sections V–VI with a detailed discussion of experimental techniques and of possible spurious effects. The proper choice and execution of precautions designed to exclude such effects constitutes the major problem in work of this kind with impurity semi-conductors.

Measurements of the conductivity of the active oxide in vacuum, as reported in the literature, have usually been incidental to studies of electron

[8] Actually, the electrochemical potential, $\bar{\mu}$, should be used throughout, where $\bar{\mu}=\mu-e\Phi_s$, $\Phi_s$ being the average electrostatic potential inside the crystal. The energy zero has been chosen so as to make $\bar{\mu}=\mu$ for the sake of simplicity.

[9] W. Heinze and S. Wagener, Zeits. f. Physik **110**, 164 (1938).
[10] W. Heinze and S. Wagener, Zeits. f. tech. Physik **20**, 16 (1939).

emission. This has generally required that conduction electrodes be separated by small and uncertain amounts. The possibilities of error caused by electrode contact resistance or polarization, or by thermionic emission through the porous oxide have not all been considered. The specimens reported in the literature have not been prepared for measurement by well-defined chemical treatment; e.g., the usual procedure has been to activate the (Ba, Sr)O for electron emission by the conventional but chemically ambiguous method of drawing current through the specimen.

The geometrical arrangement of the test specimen used in the present work and its leads is shown in Fig. 2. A chemically inert insulating support for the oxide specimen was provided by a magnesia ceramic, heated internally by a helical alumina-coated tungsten heater, as shown. The electrode system consisted of four platinum leads, 0.003 inch in diameter, mounted in appropriate grooves cut in the ceramic. This assembly was then coated by spraying with a standard cathode mixture containing $BaCO_3$ and $SrCO_3$ in the ratio of 45 to 55 by weight, suspended in a 1 percent solution of nitrocellulose binder in amyl acetate. The mean thickness of the porous carbonate specimen after evaporation of the amyl acetate was about 0.0035 inch. This assembly was sealed to a conventional high vacuum mercury pump station which maintained a pressure of $<1\times10^{-6}$-mm Hg throughout the subsequent experiments, except when their execution required treatment with purified gas.

The carbonates deposited on the magnesia were converted in vacuum to the oxides by raising the pyrometer temperature of the specimen to 1270°K over the course of about 30 minutes. The temperature was held at this level until the pressure dropped to below $10^{-6}$-mm Hg, which usually required about 5 minutes. During this decomposition the maximum true temperature was from 1400 to 1475°K. This is higher than any temperature reached during the subsequent measuring or processing of the specimen.

The (Ba, Sr)O was then heat-treated at about 1370°K until the conductivity no longer changed appreciably with time. Effective conductivity was next measured as a function of temperature, with results as shown in the lower curve of Fig. 3. The data for one tube only are shown in the inactive state, for the sake of simplicity in the diagram; the curves in the inactive states for the other tubes, whose activated states only are shown, are of the same general nature. These data in the inactive state reflect partial activation of the conductivity due to uncontrolled factors such as traces of carbon introduced as nitrocellulose, and by comparison with each other are obviously not characteristic of a well-defined or reproducible state of the (Ba, Sr)O. The measurements were made to establish a basis

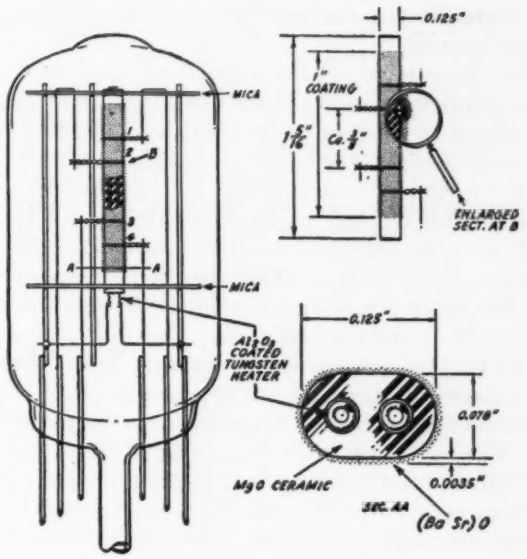


FIG. 2. Experimental tube for the measurement of conductivity.

for comparison with the results after methane treatment.

Unless otherwise specified, all temperatures reported are pyrometer readings corrected in accordance with the spectral emissivity of the specimen as described below in Section V. The primary measurement of effective conductivity was by a null method, also described below. The term "effective conductivity" is employed to emphasize the fact that electrical measurements were converted

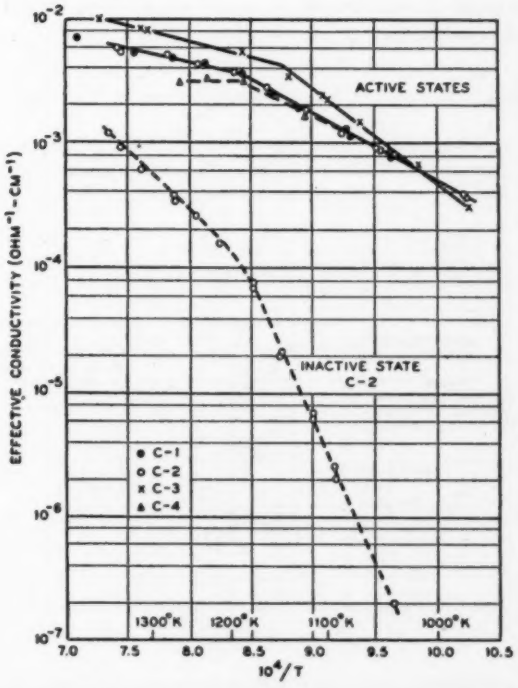


FIG. 3. Conductivity as a function of temperature in (Ba, Sr)O.

to conductivities without allowing for the porosity of the specimen.

The activation of the (Ba, Sr)O as a semi-conductor consisted of heat treatment in an atmosphere of methane. This gas was admitted to the system at room temperature until a pressure of $1.0-2.5 \times 10^{-2}$-mm Hg was reached. The specimen was then quickly heated to approximately 1320°K and held at this temperature for between 5 and 10 minutes. It was next cooled to room temperature, and the system was again exhausted. The methane treatment generally increased the spectral emissivity of the specimen, for example from 0.15 to 0.25, possibly because of deposition of invisible amounts of carbon by decomposition of the methane. The oxide was then heat-treated in vacuum at the highest temperature of subsequent measurement until spectral emissivity declined to a steady value, which required about one hour. After this final treatment the (Ba, Sr)O was ready for measurement in its activated state.

The effective conductivity of the porous (Ba, Sr)O after treatment with methane is shown in the upper curves of Fig. 3. Clearly this conductivity is everywhere larger after treatment, by a factor ranging as high as 3000. Furthermore, the increase coincided with the initiation of the methane treatment. Crude resistance measurements made during activation showed that the entire change occurred during the first two minutes, although the heat treatment in the gas continued for 5 or 10 minutes. Thus it appears that the conductivity of (Ba, Sr)O is, in general, increased by treatment with methane, as would be expected if it is an excess electronic semi-conductor.

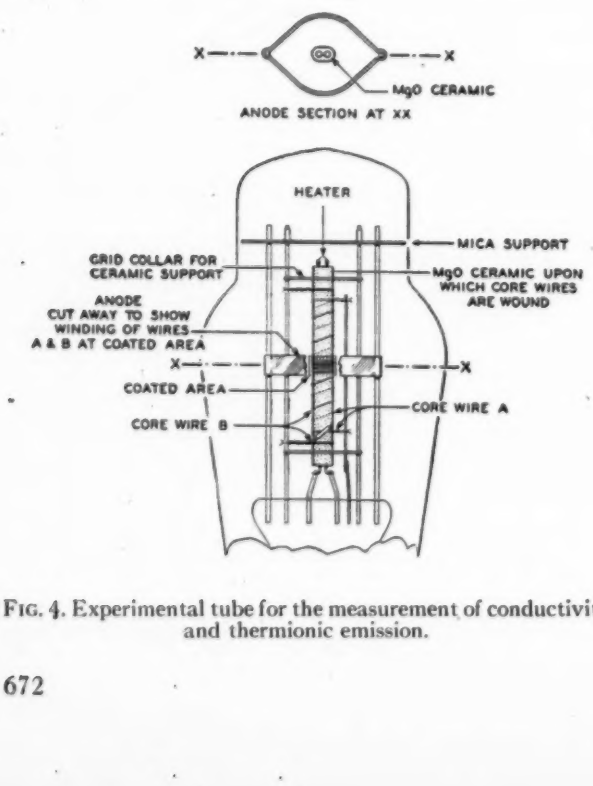


FIG. 4. Experimental tube for the measurement of conductivity and thermionic emission.

The rapidity with which steady-state values of the conductivity were obtained on initial activation is taken to be evidence that, at the temperature of the reaction, the concentration of impurity centers quickly reached a uniform value throughout the solid phase. This was probably due to the porosity of the specimen, no part of which was likely to be more than two or three microns from some pore surface. Repeated methane treatment did not increase conductivity appreciably over that observed after the initial activation. After pumping off the methane, heating at 1350°K for one hour did not change conductivity by more than 5 percent. The stability of this conductivity over the four or five hours required to complete a temperature run was repeatedly demonstrated by a final repetition of the initial measurement.

The reproducibility of the results obtained from different activated specimens as illustrated in Fig. 3, shows that at no temperature do the measurements differ from their geometric mean by more than ±50 percent. This is rather remarkable agreement for semi-conductors, particularly where the geometry of the specimens is so ill-defined, and suggests that a reasonably reproducible state of the system was attained.

The most nearly comparable conductivity values in the literature are those of Meyer and Schmidt[2] for BaO, and those of Nishibori and Kawamura,[2] apparently also for BaO. Although they were obtained from BaO which had been activated by the flow of current through the specimen (as well as by exposure to barium vapor by Nishibori and Kawamura), there is a striking general resemblance of our $CH_4$-activation results with (Ba, Sr)O to those Meyer and Schmidt at high temperatures and to those of Nishibori and Kawamura at low temperatures. In both of these experiments measurements were made with nickel electrodes a fraction of a millimeter apart.

The data recorded in Fig. 3 are in each case rather well represented by two straight lines with the lower slope at the higher temperatures. Similar behavior of conductivity as a function of temperature has been observed in ZnO,[11] and in incompletely activated BaO (Meyer and Schmidt).[2]

The slopes characteristic of the respective linear portions of the conductivity curves in Fig. 3 are 1.1 ev in the lower temperature range, and 0.3–0.45 in the higher temperature range, these slopes being expressed as values (in electron volts) of the constant $b$ in the equation

$$\sigma = Ae^{-(b/kT)}. \tag{7}$$

[11] W. Jander and W. Stamm, Zeits. f. anorg. allgem. Chemie **199**, 165 (1931); H. H. von Baumbach and C. Wagner, Zeits. f. physik. Chemie (*B*) **22**, 199 (1933).

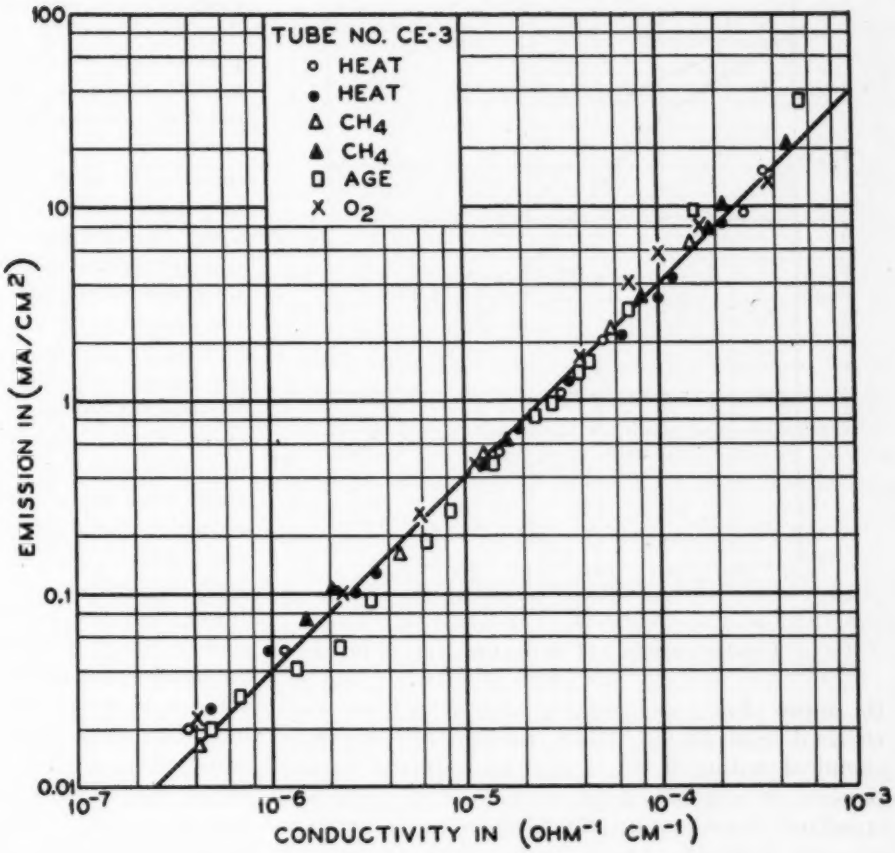


FIG. 5. Conductivity and emission upon activation of the (Ba, Sr)O, for a typical tube.

## Conclusions

These experiments justify the conclusion that (Ba, Sr)O is a "reduction" semi-conductor, i.e., it is activated as a semi-conductor by chemical treatment which tends to reduce the oxide to the free metal. The simplest hypothesis consistent with this result is that activated (Ba, Sr)O is an electronic semi-conductor because of a stoichiometric excess of (Ba, Sr) atoms in the crystal lattice.

## IV. CONDUCTIVITY-EMISSION MEASUREMENTS

The experimental diode used in this work is shown in Fig. 4. The support was the same MgO ceramic used in the conductivity work, with the cathode core consisting of two parallel platinum wires, 0.003 inch in diameter, wound on the ceramic in the form of a double helix imbedded in the oxide. The spacing of these in the central region was 0.01 inch, center to center, this region 0.10 inch in length was that covered by the (Ba, Sr)O mixture. The wires were thus separated at all points by 0.007 inch of the oxide. The anode was a bright nickel ribbon 0.2 inch wide surrounding the coated area of the cathode. The platinum core wire and the carbonate used were the same as in the conductivity work.

The tubes were left on the vacuum system throughout the course of measurement. Starting with an inactive tester, the coating freshly broken down from the carbonates, activation was carried out in slow stages. Three methods were employed: heating alone, aging with space-charge limited d.c. current, and treatment with methane at $2\times10^{-2}$-mm Hg pressure. The purpose of this was to compare

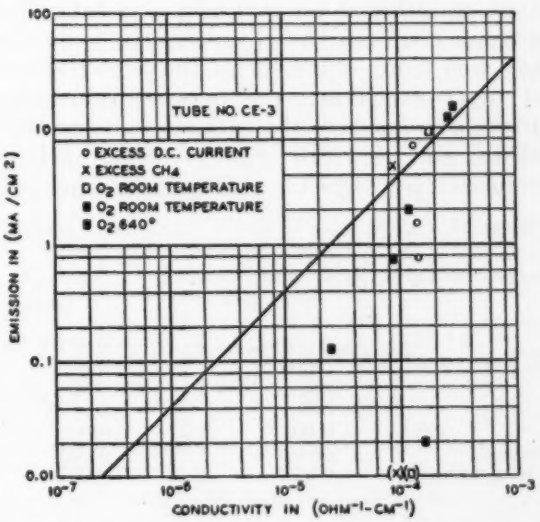


FIG. 6. Conductivity and emission showing deviations from linearity due to inhomogeneity in the oxide.

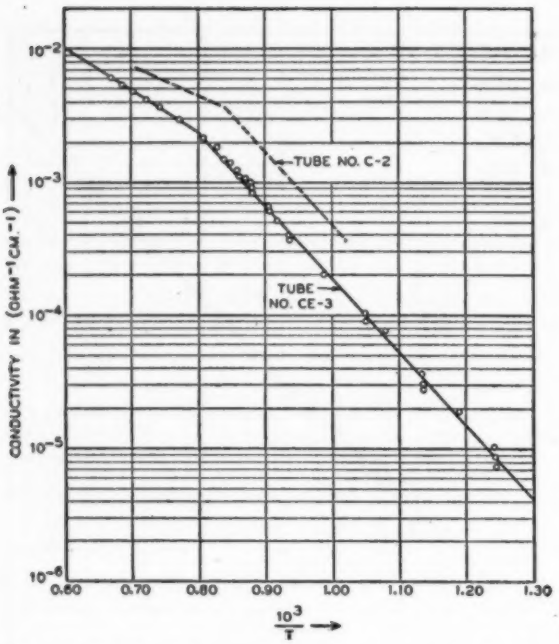


FIG. 7. Conductivity *vs.* $1/T$ in the two types of tubes.

the more clearly defined methane activation with thermal activation, which presumably involves chemical reduction by traces of reducing agents present in contact with coating, and with the standard procedure usually used in the activation of oxide cathodes, which involves the drawing of d.c. emission current. In each case, the conductivity and emission were measured after each change in state of activation, the measurements being carried out in vacuum, with the coating temperature 700°C, corrected (630°C apparent pyrometer temperature),* a temperature that was found to give no activation during the course of the measurements. Thus, all the points on the conductivity-emission curves, as in Fig. 5 for example, were taken under the same conditions, with successive points on a single line being obtained by the activation steps just described. When the maximum measurable activity was reached, the cathode was deactivated with oxygen, at $2\times10^{-2}$-mm Hg pressure for 1–2 minutes, at 700–800°C. The activation procedure could then be carried out as before. This activation-deactivation cycle was completely reversible, as the points show, no difficulty being experienced with obtaining reproducible activation curves after several oxygen treatments. The initial activation was similar to subsequent activations except that the initial activation usually proceeded somewhat more slowly. The final level of activity was about the same in each case. In the activation by drawing d.c. current, a current was chosen that amounted to about one-half that at which departure from the space-charge line set in, and a temperature was chosen at which activation by heating alone was considerably slower.

TABLE I. Quantitative analysis of carbonates used for oxide specimens.

| | $BaCO_3$ | $SrCO_3$ |
|---|---|---|
| Fe | 0.0025% by wt. | 0.001% by wt. |
| Cl | 0.006 | <0.001 |
| Ca | 0.05 | 0.20 |
| $SO_4$ | <0.01 | 0.0024 |
| Na | <0.001 | 0.02 |
| S (sulfide) | 0.001 | <0.001 |
| Acetic acid insol. | 0.013 | 0.007 |

* See Section V for comments on the temperature measurement.

Figure 5 summarizes the results obtained in the comparison of conductivity and thermionic emission over a wide range of states of activation for one tube. Five tubes were tested, the data for all being entirely similar to those in Fig. 5. Each set of points corresponds to a single activation cycle, several such sets appearing in the figure. In no case were fewer than six points obtained for an activation, and as many as twenty were sometimes obtained. In all cases the points deviate from the 45° reference line by no more than the experimental error in the emission measurements, which is the principal cause of the scatter in the points. Two heat activations, one d.c. aging activation, and two $CH_4$ activations are shown. In addition, a deactivation with $O_2$ is shown. This was obtained by starting with a completely activated tube, and proceeding with progressive deactivations by treatment with $O_2$, at 860°C, for about 2–10 seconds, at $10^{-3}$-mm–$10^{-5}$-mm Hg pressure. It will be observed that excellent agreement with the other points was obtained. Other data for this tube are shown in Figs. 6 and 7. Entirely similar data were obtained on the other tubes.

Comparison of the data for the various tubes shows that not only are the activation lines for each tube self-consistent with a slope experimentally indistinguishable from 45° in all cases, but also that the various sets of lines are nearly superimposable, the difference between the sets for different tubes being no more than the scatter among the individual lines for one tube. This indicates a remarkable degree of uniformity from tube to tube in such matters as conditions of the oxide surface, as any changes would shift the lines along one axis.

### Deviations from Linearity

The discussion of Section II implies that the oxide semi-conductor be homogeneous throughout activation, if a linear conductivity-emission relationship is to be expected. Chemical treatments, then, must affect the interior of the oxide crystals in the same way as the surface layers for this to

hold. Herring[12] has discussed the theoretical aspects of thermionic emission, concluding that usually the emission is limited by the thermodynamic state of the outer few thousand angstroms of the oxide, at most. If activation were to change only these surface layers, or the electron affinity due to a change in the surface dipole, a linear conductivity-emission relationship would not result. If, on the other hand, activation changes only the chemical potential $\mu$ in the whole body of the oxide, leaving the electron affinity $\chi$ unchanged, the linear relationship would hold.

Under certain conditions, it was found possible to obtain results that showed deviations from linearity, where this sort of inhomogeneity in the body of the oxide would be expected. These experiments are shown in Fig. 6.

The first experiment involved the deposition of carbon on the oxide particle surfaces. The data in Fig. 6 show the initial state of activity near the 45° line and the final state after deposition of carbon by treatment with excess $CH_4$ at a relatively high temperature. This is a process which has changed the outer layers but not the interior of the oxide, directly contrary to the homogeneous changes in the experiments of the last section, and results in a decrease of emission without a corresponding decrease in conductivity. The conduction paths for the electrons through the solid were essentially unaltered, whereas the surfaces from which thermionic emission occurred were completely altered by the visible deposit of carbon.

Figure 5 showed an oxygen deactivation curve which followed the same 45° straight line obtained in the activation curves. This was obtained by deactivating with oxygen at 860°C. In Fig. 6 are shown data representing two cases of deactivation with oxygen at room temperature, and one deactivation at 640°C. The treatment was for 1–2 minutes, at an $O_2$ pressure of $2\times10^{-2}$-mm Hg. In each case of room temperature deactivation, the emission dropped sharply upon $O_2$ treatment from the initial point shown near the 45° line without a corresponding decrease in conductivity. As the rates of diffusion at room temperature are undoubtedly very low, it is probable that the oxygen deactivated only the surface layers. As the temperature was raised to 700°C for measurement, diffusion would set in, of course, but not too rapidly for a measurement of the conductivity and emission before appreciable changes had occurred. The lowering of the chemical potential near the surface would serve only to decrease the thermionic emission, without greatly changing the bulk conductivity. At 640°C, the rates of diffusion were still not high enough to give a completely homogeneous coating, apparently.

In the final experiment shown in Fig. 6, a d.c. current was drawn from a fully activated tester with the plate voltage considerably in excess of that necessary to draw saturation current. Under these conditions a rapid deactivation, with regard to thermionic emission, set in, in line with general experience for these conditions, and as Fig. 6 shows, without a corresponding decrease of conductivity. It cannot be stated definitely what caused the deactivation of the surface.

[12] C. Herring, private communication.

TABLE II. Mass spectrometer analyses of purified gases.

| | | | |
|---|---|---|---|
| $CH_4$ | 97% | $O_2$ | 99% |
| Other hydrocarbons | 0.7 | A | 0.8 |
| $N_2$ | 0.66 | | |
| $H_2$ | 1.38 | | |
| $O_2$ | none detected (<0.005%) | | |

### Conductivity—Temperature Relationship

The tube used for the measurements in Figs. 5 and 6 was examined in its maximum state of activity over a considerable temperature range, to see whether the log conductivity *vs.* $1/T$ curve was similar to those obtained in the conductivity work, the data being shown in Fig. 7 along with a typical curve obtained in the conductivity work. In view of the uncertainties in the geometrical considerations in calculating conductivities from conductances, and in the relative degree of activation, it is certainly demonstrated that the two curves are measurements of essentially the same phenomena. Unfortunately, the break in this curve occurs as a temperature at which the emission measurements fail because of sparking. The curve cannot be obtained in a lower state of activation, where the emission measurements would be valid, as the tube would activate quite rapidly at temperatures approaching that of the break in the conductivity curve. Thus the possibility of a similar break in the emission curve could not be investigated.

### Conclusions of Conductivity—Emission Studies

It has been found that the electrical conductivity and the thermionic emission of a (Ba, Sr)O cathode

TABLE III. Spectrochemical analysis for contamination of (Ba, Sr)O.

| | Major (>1%) | Minor (0.1–3%) | Impurity (0.01–0.3%) | Trace (<0.03%) |
|---|---|---|---|---|
| Magnesia ceramic as received | Mg | Si | Al, Ba, Ca, Cr, Fe, Mn, Sr | Ag, Cu, Mo, Na, Ni, Ti |
| Carbonate spray mixture | Ba, Sr | Ca | Na | Al, Bi, Cu Fe, Mg, Mo, Pb |
| (Ba, Sr)O after conductivity tests | Ba, Sr | Ca | Fe, Mg, Na | Al, Cu, Mn, Ni, Si |

with a Pt core are directly proportional through three orders of magnitude of activation: the activation was accomplished by heating without drawing current, by heating while drawing space-charge limited d.c. current, and by reduction with methane gas, and a deactivation curve was obtained with oxygen at 860°C. In all cases the same linear result was obtained, regardless of the method of activation or deactivation; furthermore, on several experimental tubes exactly the same results were found, no appreciable shift of the lines from tube to tube being observed. Thus a quantitative relationship involving the thermionic emission, predicted by the semi-conductor theory for the activation process, has been experimentally verified within the limits of error of measurement. For a cathode of this particular geometry, if one is assured that equilibrium conditions prevail, a knowledge of the electrical conductivity enables a prediction of the thremionic emission current to be made. It may be concluded that activation changes only the chemical potential, little or no change in the electron affinity, $\chi$, occurring. This experimental result is inconsistent with theories of activation involving changes in free barium on the surface of the oxide, such as have been proposed in the past. Moreover, this proportionality of conductivity and emission would not be expected if there were an appreciable number of surface states on the oxide.

In addition, it has been found that certain processes which would be expected to change only the surface layers of the oxide will affect the emission without a corresponding change in conductivity, these processes being (1) deposition of carbon on the crystal surfaces with excess $CH_4$ treatment; (2) oxygen treatment at room temperature, with inhomogeneity arising from the low diffusion rates at this temperature; and (3) processes arising from the drawing of saturation d.c. current. These observations are consistent with the view that thermionic emission is determined by the thermodynamic state of the outer atomic layers.

## V. DETAILS OF EXPERIMENTAL TECHNIQUES

Some of the precautions taken to insure the validity of these results will be described in this and the next section.

TABLE IV. Resume of spectral emissivity measurements on (Ba, Sr)O specimens.

| Specimen | | Number of measurements | Temperature range | $E_\lambda(\lambda=0.66\mu)$ |
|---|---|---|---|---|
| C-2 | before activation | 11 | 300°–1350°K | 0.17 ±0.015 |
| | after activation | 8 | 300°–1270° | 0.21 ±0.025 |
| C-1 | before activation | 6 | 300°–1340° | 0.10 ±0.015 |
| | after activation | 9 | 300°–1320° | 0.09 ±0.025 |
| C-3 | before activation | 7 | 990°–1350° | 0.14 ±0.025 |
| | after activation | 4 | 980°–1140° | 0.15 ±0.015 |

## Chemistry of Materials and Processes

The principal impurities found in the original carbonates are shown in Table I.

The gases used in the course of investigation, methane and oxygen, were purified before admission to the vacuum system, their purity being checked by mass spectrometer analyses. The methane was passed over hot copper and anhydrous calcium sulfate, and also over Ascarite in the case of the conductivity-emission work. The oxygen was passed over Ascarite and anhydrous calcium sulfate. In each case the liquid-nitrogen trap remained on the system. Typical analyses after purification are shown in Table II.

The methane used in the conductivity experiments also contained 0.25 percent $CO_2$, but this was removed by condensation in a liquid-nitrogen trap before admission to the (Ba, Sr)O specimen.

One of the chief problems was to select for the insulating support of Figs. 2 and 4 a material which would not react while in contact with (Ba, Sr)O at high temperatures. Magnesium oxide fired in an oxidizing atmosphere was chosen because of its failure to form any known compounds with alkaline earth oxides.[13] X-ray and electron diffraction data obtained by Mr. K. H. Storks of these laboratories showed no evidence of reaction between BaO and MgO after heating the powdered mixture to 1000°C in hydrogen. At the conclusion of one of the conductivity experiments, some of the (Ba, Sr)O was scraped off its support and compared by qualitative spectrochemical analysis with the original carbonate suspension. The results are recorded in Table III. It is apparent that the ceramic did not react appreciably with the (Ba, Sr)O and indeed contaminated it only slightly. The only new impurities (0.01–0.3 percent) are iron and magnesium, presumably as oxides from the ceramic. In these quantities they cannot affect the result of methane treatment except for the remote possibility that (Ba, Sr)O is a semi-conductor by virtue of unoxidized iron or magnesium dissolved in it.

It was also essential to determine the effect of heat treatment in methane on the insulation resistance of the magnesia support. After successive treatments in the absence of (Ba, Sr)O at 1400°K in vacuum and 1300°K in methane, this resistance between leads 2 and 3 of Fig. 2 remained too high ($>10^9$ ohms at 1400°K) for convenient measurement.

Platinum was chosen for the metallic leads to the specimen because of its chemical inertness, low vapor pressure, and spectroscopic purity, except for a trace (<0.03 percent) of copper.[14]

[13] G. A. Rankin and H. E. Merwin, J. Am. Chem. Soc. **38**, 568 (1916). They report, for example, that MgO and CaO do not form compounds.

[14] This platinum was prepared by Mr. H. T. Reeve of these laboratories.

Another chemical question is the effectiveness of methane in reducing (Ba, Sr)O under the specified conditions. Such reduction is shown to be thermodynamically reasonable by estimating the Gibbs free energy at 1273°K for the reaction,

$$BaO(solid)+CH_4=Ba(gas)+CO+2H_2.$$

This free energy turns out to be +55,000 calories.[15] The corresponding equilibrium constant $K$ is $4\times10^{-10}$; furthermore $K=(p_{Ba}\cdot p_{CO}\cdot p_{H_2}{}^2)/p_{CH_4}$. This together with our idealized conditions that $p_{Ba}=p_{CO}=\frac{1}{2}p_{H_2}$ yields $p_{Ba}=3\times10^{-3}\cdot p_{CH_4}{}^{\frac{1}{3}}$, where the partial pressures $p$ are identified by the chemical symbols in the subscripts and are expressed in atmospheres. The equilibrium pressure of barium vapor in the presence of 0.025-mm Hg of methane is thus calculated to be 0.2 mm, with an accuracy of about a factor of ten. This is about one-tenth of the vapor pressure of solid barium as extrapolated to 1273°K from Van Liempt's data.[15] Hence equilibrium conditions should be favorable to the production of gaseous barium and its solution by solid BaO to the extent of 0.01 or more of its solubility limit at this temperature. Schriel's work suggests a solubility limit of about one atom percent of Ba in BaO at about 1420°K.[16]

A mass spectrometer analysis of the gaseous products of the reaction showed complete recovery as CO of the carbon in the $CH_4$ (with small amounts of $CO_2$ and unreacted $CH_4$), as expected from the above. However, a control tester containing no (Ba, Sr)O yielded a similar complete recovery of methane carbon as CO, showing that other hot oxides such as MgO, $Al_2O_3$, or $WO_2$ compete with (Ba, Sr)O in reacting with methane.

Another possibility is that methane treatment produced a second solid phase of conducting material deposited on or scattered through the (Ba, Sr)O. The complete recovery of the carbon in the mass spectrometer experiment indicates that no considerable deposit of carbon was formed. This is confirmed in Table IV, which shows the spectral emissivity measured concurrently with conductivity as discussed below. Methane treatment produced no net change of emissivity greater than experimental error. Small quantities of carbon or other good conductors would be expected to affect this property. Hence it is concluded that no second conducting phase is produced by the methane treatment.

The chemical evidence then supports the conclusions that in these experiments reasonably pure (Ba, Sr)O is available for reaction with similarly pure methane; that the significant effect of the methane is to reduce small quantities of (Ba, Sr)O, and that after the reaction the conducting specimen consists of (Ba, Sr)O which should contain at least a fraction of any excess barium that it is capable of dissolving.

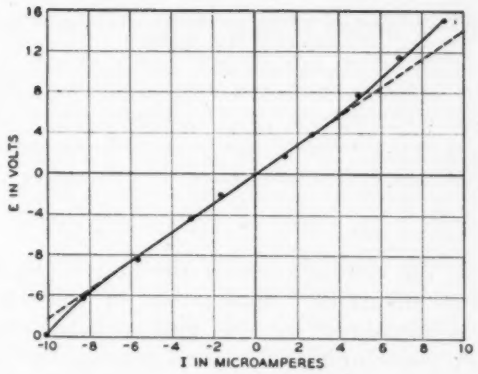


FIG. 8. Probe voltage *vs.* current in (Ba, Sr)O before activation, at 1176°K. Tube C-2.

### Measurement of Conductivity and Temperature in Conductivity Experiments

The primary measurement of effective conductivity of the porous (Ba, Sr)O specimen was the measurement of the potential difference between probes 2 and 3 (Fig. 2) as a function of the current through an activated specimen, at several temperatures. The behavior of such specimens was invariably ohmic over the entire range of observations, which included current densities as high as 65 ma/cm². The resistance determined by a slope of such current-voltage measurements was combined with the dimensions of the specimen as originally sprayed to calculate effective conductivity. For this purpose the area of cross section was determined by the weight of the coating, together with its length and its bulk density. This bulk density was found to be 1.1 g/cm², as measured after depositing a test specimen according to the same spraying procedure on the flat metal block. These dimensions take no account of over-all shrinkage of the specimen during decomposition of the carbonates, nor of its obvious porosity. Nevertheless the data are reported as effective conductivity to facilitate comparison among the results recorded here and elsewhere in the literature.

The conductivity was observed at successively increasing temperatures from about 1000°K to 1350°K, with a final check at the initial temperature to detect any irreversible change occurring during the four hours involved. After activation there was no such change of more than 10 percent.

[15] The estimate is based on data from K. K. Kelly, U. S. Bur. Mines Bull. No. 371 (1934); No. 383 (1935); No. 394 (1936); together with the following assumptions:

(1) The high-temperature molar heat capacity of BaO equals that of CaO.
(2) The heat of sublimation of barium at 298°K is +44,500 calories which is the value observed at 800°K–1050°K by J. A. M. Van Liempt, Rec. Trav. Chim. de Pays-Bas **55**, 468 (1936).

[16] M. Schriel, Zeits. f. anorg. allgem. Chemie **231**, 313 (1937).

TABLE V. Effect of helium on effective conductivity of activated (Ba, Sr)O.

| Specimen | $T$ | $\sigma$ In vacuum | $\sigma$ In He 1 atmos. | $\sigma$ Re-evacuated |
|---|---|---|---|---|
| C-3 | 1303°K | $8.1\times10^{-3}$ ohm$^{-1}$ cm$^{-1}$ | $8.2\times10^{-3}$ ohm$^{-1}$ cm$^{-1}$ | $8.2\times10^{-3}$ ohm$^{-1}$ cm$^{-1}$ |
| C-2 | 1304°K | $3.9\times10^{-3}$ | $4.1\times10^{-3}$ | $4.7\times10^{-3}$ |
| C-2 | 1066°K | $8.6\times10^{-4}$ | $7.3\times10^{-4}$ | $7.3\times10^{-4}$ |
| C-1 | 1247°K | $4.9\times10^{-3}$ | $2.5\times10^{-3}$ | $3.8\times10^{-4}$ |

While the conductivity of the active (Ba, Sr)O was thus invariably ohmic, that of the inactive state as measured by the null method was invariably non-ohmic in the sense typified by Fig. 8. For BaO the current-voltage curves of Meyer and Schmidt[2] exhibit the same general behavior in the active and inactive states respectively. This tendency towards saturation of the current in the inactive state may reflect thermionic emission in addition to conduction current, as discussed in Section VI. The slope $(dE/dI)_{I=0}$ was taken to be the resistance in determining conductivity as a function of temperature as shown in the lowest curve of Fig. 3. This procedure is justified insofar as the results are used to show that methane treatment increases conductivity.

An alternative method of measuring the resistance of the specimen lying between the probes 2 and 3 of Fig. 2 was to connect these probes to a d.c. Wheatstone bridge. The results were found to agree closely with those of the null method for the activated state and to disagree extensively in the case of only one measurement of the inactive state. Hence the contribution of electrode contact and polarization effects is negligible in this experimental arrangement of a two-electrode system.

Temperatures were measured with calibrated optical pyrometers. In order to correct these pyrometer readings in accordance with the spectral emissivity of the specimen, a reflectometer of the type developed by Prescott and Morrison[17] was installed. The use of this technique with the translucent (Ba, Sr)O, supported by a translucent ceramic containing a heater, raises several questions. The method relies on the determination of the spectral emissivity $E_\lambda$ from the measured reflectivity $R_\lambda$ according to

$$E_\lambda + R_\lambda = 1. \tag{8}$$

This in turn requires that only a negligible fraction of the light incident on the tester should be transmitted through it.[18] A rough measurement of this fraction for the oxide-coated magnesia yielded a value of $<10^{-3}$, justifying the use of Eq. (8).

The spectral emissivity was measured concurrently with the conductivity throughout these experiments. The values obtained at the wavelength of $0.66\mu$ are recorded in Table IV for three specimens. In all cases these results were independent of temperature within the precision of the measurement; hence the table records mean values and their mean deviations over the indicated temperature ranges. The variation of $E_\lambda$ among these three specimens may reflect differences of granularity. The mean values of Table IV were used to determine the true temperature $T$ from the pyrometer reading $D$ in the standard way, according to

$$1/T - 1/D = (\lambda/c_2)\ \ln E_\lambda, \tag{9}$$

where $c_2$ was taken to be 14,380 micron-degrees. These corrections were so large that the uncertainty in $T$ associated with that in $E_\lambda$ is of the order of $\pm 25°$.

The correctness of these temperature measurements involves the assumption that the light employed in temperature measurement originates primarily in the (Ba, Sr)O rather than in the translucent supporting ceramic or the tungsten heater. This was tested by several measurements of the conductivity of activated (Ba, Sr)O surrounded by helium at a pressure of one atmosphere; such measurements were compared with observations at the same pyrometer temperature in vacuum. The results in Table V show remarkable agreement except for specimen C-1, where a partially irreversible change is ascribed to chemical effects.

If the assumption concerning the origin of the light is valid, the presence of an atmosphere of helium should make no difference, but if it is not, the thermal conductivity of the helium should act to increase the true temperature, and thus the electrical conductivity of (Ba, Sr)O compared with that in vacuum for a given pyrometer reading. Thus in a typical case at a pyrometer temperature of 1190°K ($T=1304$°K), the heater temperature as estimated from its resistance was 1540°K in vacuum and 1350° in helium. The intensity of $0.66\mu$ light received from the heater and ceramic was certainly less in helium than in vacuum. Hence if our assumption was invalid, more light was received from the (Ba, Sr)O in helium than in vacuum, at this pyrometer reading. This would then require that the temperature and conductivity of (Ba, Sr)O be higher at a given pyrometer reading in helium than in vacuum. The observed agreement thus indicated that within the limits of error of the experiments and their interpretation, the primary origin of the light lay in the oxide, except for the possibility of error in the above data due to an improbable compensation of spurious effects. For example, if a large portion of the current were due to thermionic emission around the specimen or through its pores,

[17] C. H. Prescott, Jr., and James Morrison, Rev. Sci. Inst. 10, 36 (1939).

[18] American Institute of Physics, *Temperature, Its Measurement and Control in Science and Industry* (Reinhold Publishing Corporation, New York, 1941), Chapter 1.

such compensation is conceivable. This latter effect will be considered in detail in Section VI, as it is of fundamental importance in all of the present work.

### Techniques in Conductivity—Emission Measurements

After winding the platinum wires on the MgO ceramic, the structure was cleaned with solvents, vacuum and hydrogen-treated, and sprayed with the carbonate mixture to a thickness of 0.005 inch over the ceramic. Care was taken during the spraying process to pack the carbonate down well between the individual wires, the first part of the spraying as "dry" as possible, with the final layers on top of the wires being applied in the usual "wet" spray.

The tube, when mounted on the vacuum system, was given a preliminary 2-hour bake-out at 400°C. The anode and support collars were then outgassed at 900°C, by high frequency induction heating. The carbonates were broken down thermally, at temperatures up to 900°C, until the pressure in the system dropped below $5\times10^{-7}$-mm Hg.

Temperatures were measured with an optical pyrometer, although once the relationship between pyrometer temperature and heater power was established for a given tube, the latter was taken as the more reliably reproducible indication of temperature. Corrections to true temperature were made by using the spectral emissivity data obtained in the conductivity work, the emissivity being taken as 0.15. Due to the difficulty of investigating with certainty the possibility of light originating in the heater-ceramic system, it cannot be stated definitely that the true temperature of the coating was not below that reported, however.

The conductivities were measured using the two interlacing platinum wires as a two-electrode system. Justification for this lay in the results of Fig. 7, and in the conductivity work, where it was found that 4-electrode potentiometric measurements agreed with 2-electrode bridge measurements in every case where platinum electrodes were used, for that geometry.

The conductivities were all obtained from the slopes of current-voltage curves taken between the two leads, several points being taken on both sides of the origin at a low enough voltage gradient to give ohmic results. It was observed that non-ohmic behavior of the type shown in Fig. 8 usually set in at about 0.5–1.5 volts between the leads; the slopes were all taken from linear current-voltage measurements in the region up to 0.05-v maximum between the leads in the higher activity ranges, 0.1–0.2-v in the low activity ranges. The observed conductances were converted into apparent conductivities for the porous oxide using the best available geometrical data for the cathode structure.

Thermionic emission was measured by a pulse technique, single 1/120-second sine-shaped continuously variable voltage pulses being applied between the plate and the two core wires, which were connected together externally for this measurement. A signal proportional to the current through the tube was applied to the vertical axis of a cathode-ray oscilloscope, and a signal proportional to the three-halves power of the voltage developed across the tube was applied simultaneously to the horizontal axis. Thus, the oscilloscope trace gave a complete 3/2 power plot, which provided a check at all times as to possible fading, sparking, and gas effects, with a straight line oscilloscope trace as long as the emission was space-charge limited. The point at which the current fell 20 percent below the space-charge line was taken as a measure of the emission as shown in Fig. 9, the method being largely arbitrary, as are most emission measurements. The voltage was increased in each case above that just necessary to give 20 percent deviation, to observe the effect on the current at the 20 percent deviation point. This usually increased slightly as the voltage rose, approaching a more or less constant value, which was the current actually used as a measure of the activity of the cathode. The crossing of the 20 percent line took place after about the first two milliseconds of the 1/120 second pulse, several single pulses being applied within about a minute to obtain the measurement.

## VI. MECHANISM OF CHARGE TRANSFER

The conductivity involved in the discussion of Section II is the bulk electronic conductivity of the oxide semi-conductor, and the thermionic emission is assumed to be that of electrons emitted from the surface layers after transfer through the solid from the core metal to the surface layers. It is, therefore,

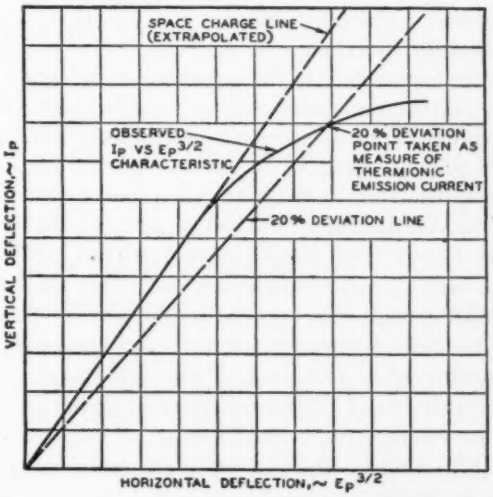
SPACE CHARGE LINE (EXTRAPOLATED)
OBSERVED $I_p$ VS $E_p^{3/2}$ CHARACTERISTIC
20 % DEVIATION POINT TAKEN AS MEASURE OF THERMIONIC EMISSION CURRENT
20% DEVIATION LINE
VERTICAL DEFLECTION ~ $I_p$
HORIZONTAL DEFLECTION, ~ $E_p^{3/2}$

FIG. 9. Plot of oscilloscope trace showing method used in measuring thermionic emission.

TABLE VI.

| | Conductivity ($ohm^{-1}$ $cm^{-1}$) | Emission ($ma/cm^2$) |
|---|---|---|
| Before $CH_4$ treatment | $8.8\times10^{-5}$ | 4.7 |
| After $CH_4$ treatment | $8.8\times10^{-5}$ | <0.01 |
| Before $O_2$ | $1.8\times10^{-4}$ | 9.5 |
| After $O_2$ (room temp. exposure) | $1.2\times10^{-4}$ | <0.01 |
| Before $O_2$ | $2.9\times10^{-4}$ | 15.0 |
| After $O_2$ (room temp. exposure) | $1.6\times10^{-4}$ | 0.02 |

a vital point that this be the actual mechanism of charge transfer. The question of the method of charge transfer through the oxide is of fundamental importance to any theory of the behavior of oxide cathodes, and is one which is treated inadequately in the experimental literature. The two most possible sources of error in conductivity to be considered are (1) that there was an ionic conductivity to a sufficient extent to obscure the results, and (2) that it was not a bulk property of the solid being measured, but rather electron transfer by thermionic emission through the pores of the material or around the material, conductivity along the pore walls, high resistance contacts between individual crystals, or a property of the metal-oxide interface.

Concerning the first of these, attempts to detect any polarization by passing a high d.c. current through the coating for long times, failed to show any detectable reverse current when the applied voltage was suddenly shut off. The conductivity behaved like excess electronic semi-conduction insofar as it was increased by reducing agents and decreased by oxidizing agents. The indication is that the ionic contribution to the conductivity was small.

Concerning the possibility that the electron transfer was not really a bulk property of the solid, it is considered unlikely that it was a property of the metal-oxide interface. In the case of the conductivity measurements, the agreement in every case where platinum electrodes were used between 4-electrode potentiometric measurements and 2-electrode bridge measurements, indicates there was no significant interface resistance. In a few experiments using Grade A nickel electrodes, on the other hand, it was found that these measurements no longer agreed after oxidation treatment, and that conductivities were not reproducible above 1000°C. The absence of appreciable interface resistance in the conductivity tester, and the agreement of the conductivity *vs.* $1/T$ curves for this structure with that for conductivity-emission structure, are considered to justify the use of 2-electrode measurements in the latter case.

The possibility that thermionic emission current in parallel or in series with conduction current in the solid played an important role must be considered in detail. The specimens studied here were porous, and exhibited fine cracks extending through at least part of the thickness of the specimen, as is usual with oxide coatings. Except for the experiments of Becker and Sears at much lower temperatures,[2] little consideration has been given to such spurious effects. A number of experiments were carried out in the present work to investigate this possibility, not all of which were conclusive. The approach has been to subject the specimen to treatments which would affect thermionic emission current but not conduction in the solid, and to look for changes in the observed conductivity during such treatment. Three methods of accomplishing this have been the introduction of an atmosphere of helium, the application of a magnetic field, and the deactivation of surface layers of the oxide crystals by proper chemical treatment.

In the helium experiments, vacuum measurements were compared with conductivities obtained in 1 atmos. of helium, the specimen being heated to the same pyrometer temperature. Since the mean free path of a thermal electron in helium can be estimated from the literature[19] to be no more than $1\times10^{-4}$ cm at 760 mm and 1000–1300°K, and since most of the volume of the bulk specimen is visibly occupied by voids of much larger dimensions, this test should detect any important spurious effect. The results of such experiments on conductivity testers in the activated state are shown in Table V. Before admission, the helium was purified with copper at 400°C, and anhydrous $CaSO_4$; mass spectrometer analysis showed only 1.4 percent $N_2$ and 0.1 percent $H_2$. The effect of the gas heating mica and glass surfaces resulting in transfer of contaminants to the specimen appears to have been small, in view of the reproducibility obtained.

It was then observed that a magnetic field of 2000 gauss reduced the current in a conductivity tester in the inactive state by as much as 50 percent, the effect being reduced after activation to the range 0–20 percent. In no case in scores of tests made on three conductivity specimens in various states of activation, at temperatures ranging from 1000–1300°K, were effects of more than 1 percent found to be produced by a magnetic field with an atmosphere of helium present. Such small changes could be expected from the effect of the field on the geometry of the tube, including the heater coil. Magnetic effects in vacuum of a similar magnitude were observed in conductivity-emission testers, also.

The third type of evidence bearing on the question is that shown in Fig. 6 and described in Sec-

[19] R. B. Brode, Rev. Mod. Phys. **5**, 257 (1933).

tion IV, in particular with regard to the experiments involving deposition of carbon, and room temperature oxygen deactivation. Three such experiments on conductivity-emission testers are shown in Table VI.

If the electron transfer through the solid were due primarily to thermionic emission currents in the pores and cracks, rather than electronic semiconduction through the solid with subsequent emission from the outer oxide surfaces in the case of thermionic emission current to the anode, the conductivity should have been decreased in the same way as the emission current. In the case of the oxygen deactivations, the small change in the conductivity may be ascribed to the fact that as the temperature was raised to 700°C for the measurements, diffusion would become more rapid, and the effect of the deactivated surface layers would begin to appear throughout the body of the oxide.

It appears established from these data that in the activated state conductivities, as measured in these specimens, reflect primarily the flow of electrons through the solid; in the inactive state spurious thermionic currents are of more importance. However, it is concluded that with an error no greater than that of the emission measurements, the conductivity measurements reflect solid conduction, in view of the following:

(1) In the activated state this current was not affected by the presence of helium at one atmosphere pressure.
(2) Processes designed to affect the surface layers throughout the porous specimens greatly reduced the emission to the anode but did not affect the conductivity.
(3) Conductivity measurements made in specimens of widely variant geometry agree within reasonable error.
(4) The current-voltage curves in the conductivity-emission testers were ohmic up to a voltage gradient of the order of 50 v/cm over three orders of magnitude in the conductivity.


## ACKNOWLEDGMENT

This work has had the benefit of the aid and advice of several members of these laboratories, in particular: Dr. J. A. Burton, Dr. C. Herring, Dr. J. T. Smith, and Dr. L. A. Wooten.

# Thermionic Emission from Sintered Cathode of Thoria and Tungsten Mixture*

H. Y. Fan**
*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts*
(Received November 15, 1948)

To obtain suitable conductivity for direct heating, tungsten powder is added to sintered thoria cathodes. The thermionic properties of a cathode, sintered from a mixture of 67 percent thoria and 33 percent tungsten, were investigated. Emission constants were determined and the emission was found to be somewhat lower than that of cathodes sintered from pure thoria.

Change of cathode activity with temperature was observed and studied in detail. The cathode showed high activity at temperatures which would bring about deactivation of ordinary thoriated tungsten and decreased its activity when the temperature was lowered to the range where thoriated tungsten would be activated. A theory is proposed for the mechanism of the change in activity, involving the production of free thorium due to reduction of thoria by tungsten, diffusion and evaporation of thorium from thoria. Evaporation of thorium as well as thoria was proved by the use of a tungsten filament probe which detected evaporation of materials from the cathode by the change of its own emission. The energy of evaporation of free thorium from thoria is estimated from the experiments to be 46600 cal./g-atom. The energy of evaporation of thoria is found to be 184000 cal./g-molecule. The behavior of the cathode is satisfactorily explained by the proposed theory.

## INTRODUCTION

SINTERED thoria cathodes, developed in recent years, possess a number of desirable properties,[1] of which the two most important are: (1) The ability to withstand severe punishment caused by sparking, bombardment and other destructive phenomena. (2) Operation at higher temperatures, as compared with ordinary oxide-coated cathodes, with consequent ability to dissipate extra heating power, at the same time quite a large density of electron emission is obtainable. These properties make this type of cathode especially suitable for applications such as high powered magnetrons. However, due to the extremely high resistivity of thoria it is necessary to provide such cathodes with a separate heater which has to be run at very high temperatures. The design of such heaters is a problem and burn out of the heater is frequently a source of trouble.

To simplify the problem of heating, cathodes are now being developed which are sintered from thoria mixed with tungsten powder.[2] By varying the percentage of added tungsten the resistivity of the cathode can be adjusted to a value suitable for direct heating. The cathode studied here consists of 67 percent thoria and 33 percent tungsten by weight. It is made in the form of a hollow cylinder pressed from such mixture and fired in a hydrogen furnace at 1800°C for one half hour. The dimensions of the cathode are: 1″ in length, 0.265″ outside diameter and 0.192″ inside diameter. Molybdenum sleeves platinum brazed to the two ends serve as current leads. Figure 1 shows the resistivity $\rho$ and the required heating power $W$ as function of temperature. The resistivity of tungsten $\rho_w$, multiplied by a factor 845 to adjust it to the same value as the resistivity of the cathode at room temperature, is shown for comparison. During the operation of over five hundred hours at temperatures varying from 1100°K to 2300°K the resistance of the cathode, as measured between the cathode leads of the tube in which the cathode is mounted, increased by approximately 10 percent. The results

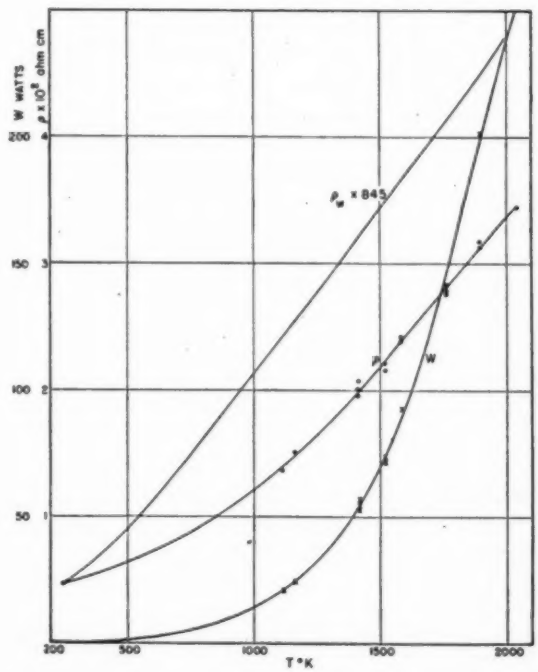


FIG. 1. Heating power and resistivity of the cathode as functions of temperature.

* This work has been supported in part by the Signal Corps, the Air Materiel Command, and ONR.
** On leave of absence from National Tsing Hua University, Peiping, China. Now at Department of Physics, Purdue University, Lafayette, Indiana.
[1] M. A. Pomerantz, N.D.R.C. Division 14, Report 517 (1945).
[2] This development seems to have been originated by the Raytheon Manufacturing Company.

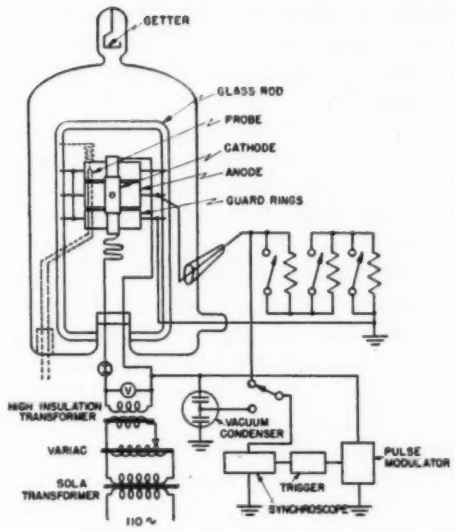


FIG. 2. Diagram of tube structure and measurement circuit.

of the study of the properties of such a cathode are reported here.

## EXPERIMENTAL ARRANGEMENT

A conventional diode structure is used for the study of thermionic emission (Fig. 2). The anode is a cylinder, concentric with the cathode, 1″ in diameter and $\frac{7}{16}$″ long. It is provided with guard rings, each of 1″ in diameter and $\frac{1}{2}$″ long. For the measurement of true temperature a hole, 10 mil in diameter, is drilled through the wall in the middle of the cathode. A larger hole is made in the anode so that the hole in the cathode wall can be viewed conveniently with an optical pyrometer. Before sealing off the pump the vacuum as read by an ionization gauge was below $10^{-7}$ mm Hg with the cathode at a high operating temperature and with a d.c. emission giving an anode dissipation several times the highest value ever reached during the operation after seal-off. The tube is also gettered. The probe arrangement shown by the dotted lines was added in a second tube only. It will be described later.

D.c. emission was measured for low temperatures of the cathode. For higher temperatures a pulse modulator was used which gives pulse lengths from 0.5 to 10 μsec. The pulse current and voltage were measured on the screen of a synchroscope. The circuit is shown schematically in Fig. 2. The cathode was heated from the a.c. line through a Sola constant voltage transformer, a variac and a high insulation transformer in series. The last mentioned was used because the cathode was raised to high negative voltages during the pulses.

## EMISSION IN ACCELERATING FIELDS

Figure 3 shows the curves of $I$ as function of $V^{\frac{1}{2}}$. It is seen that for sufficiently high fields the curves approach the theoretical Schottky slopes. Since the surface of the cathode consisting of thoria and tungsten mixtures is nonuniform, deviation from the theoretical slopes at low fields is expected on the basis of patch theory.[3] The fact that the curves approach the theoretical slopes more readily for higher temperatures is also consistent with the predictions of the patch theory. The sharp bending away from the theoretical line in the curve for 2040°K is the transition into space charge limited region; below the bend the curve follows closely the calculated $\frac{3}{2}$ power curve.

Since the cathode is directly heated its zero field emission cannot be measured readily. As the emission approaches Schottky lines for high voltages we may well use the emission values at a definite, sufficiently high voltage for the determination of emission constants. All emissions were accordingly measured at 2500 v. A still higher voltage would be preferable, as the emissions at low temperatures have not quite reached the theoretical slopes at this voltage. But we were limited in this respect by the pulse modulator. It can be seen from Fig. 3 that a higher voltage will not make a great difference.

The emission under applied d.c. voltage at low temperatures and under pulse voltage at high temperatures maintained a definite value and showed

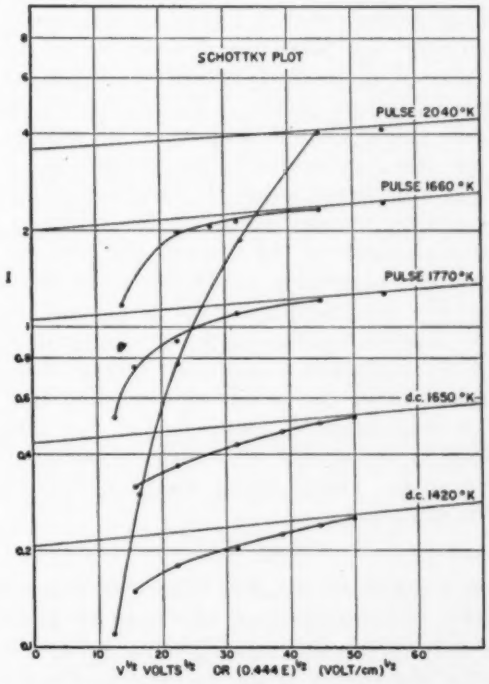


FIG. 3. Schottky plot of emission current. Current scales for different curves: 1=2.0 amp. (2040°K); 1=0.126 amp. (1770°K); 1=10 ma (1660°K); 1=10 ma (1650°K); 1=0.20 ma (1420°K); d.c. and pulsed emission curves are not for the same state of the cathode.

[3] J. A. Becker, Rev. Mod. Phys. **7**, 95 (1935).

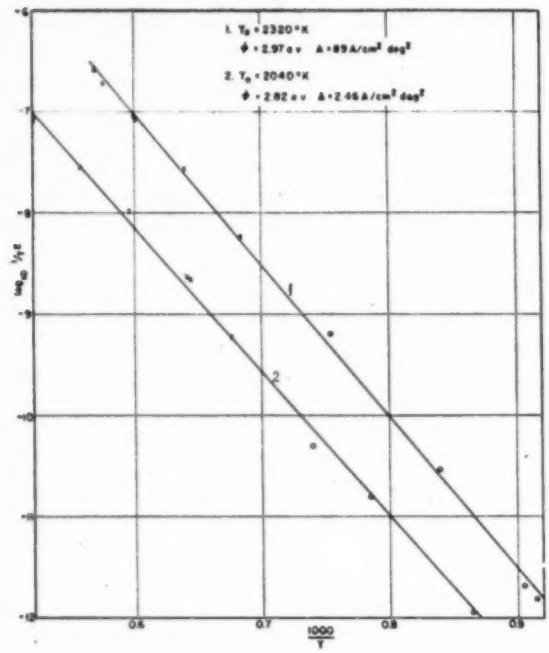

FIG. 4. Richardson plot for equilibrium state of the cathode at two different temperatures.

no sign of decay. For temperatures where both d.c. and pulsed emissions could be measured conveniently the values were the same. This result shows that there is no fundamental difference between emissions under d.c. and pulsed conditions. Furthermore the tube was opened to air and rebuilt twice. The emission properties of the cathode showed no marked change. This is indeed one of the desirable properties of this type of cathode. If however, the temperature was changed and maintained at a different value the emission underwent a time change irrespective of applied voltage. This indicates a change of the state of the cathode with temperature. This effect will be discussed more fully later on. For determining the emission constants the cathode was maintained at a definite temperature $T_0$ and the emissions at different temperature $T_t$'s were taken to be the initial values when the temperature was brought from $T_0$ to $T_t$. Thus all emissions corresponded to a definite state of the cathode at $T_0$. Thoria being a semi-conductor, its emission equation

$$i = AT^n e^{e\phi/kT} \qquad (1)$$

may have different possible values of $n$ according to theory, depending upon the type of semi-conductor it is. Without knowing what type of semi-conductor the thoria in the cathode is, there is no basis for taking any particular value for $n$. We shall use the conventional value $n=2$ for convenient comparison of the emission constants $A$ and $\varphi$ with data given by other authors. Figure 4 shows curves of $\log(i/T^2)$ *versus* $1/T$ for two values of $T_0$: 2040°K and 2320°K. Both $\varphi$ and $A$ are different for the cathode states at these two temperatures. The state at 2320°K gives higher emission on account of higher value of $A$.

Wright[4] gives for thoria coated on tantalum $\varphi=2.5(4)$ ev, $A=2.5$ amp./cm² deg.² for d.c. and $\varphi=2.6(2)$ ev, $A=7.5$ amp./cm² deg.² under pulsed conditions. He gives also $\varphi=2.5(5)$ ev, $A=3.5$ amp./cm² deg.² for the pulsed emission from an extruded thoria sleeve heated by a tungsten spiral. Hanley[5] gives for cataphoretically coated thoria cathodes average values: $\varphi=2.67$ ev, $A=2.63$ amp./cm² deg.² under d.c. conditions and $\varphi=2.55$ ev, $A=5.62$ amp./cm² deg.² under pulsed conditions. Our values of $\varphi$ are higher than those given by these authors. This result may be due to the fact that a part of the surface of our cathode is presented by tungsten which has a higher work function. Dependence of emission constants on the cathode temperature was also observed by Weinreich[6] for thoria coated on tungsten. Contrary to our results the constant $A$ remains unchanged while the work function is lower for higher temperature of the cathode.

## TEMPERATURE DEPENDENCE OF ACTIVITY

It has been mentioned that the cathode activity shows temperature dependence. In fact the emission constants have different values for different conditioning temperatures of the cathode. The observa-

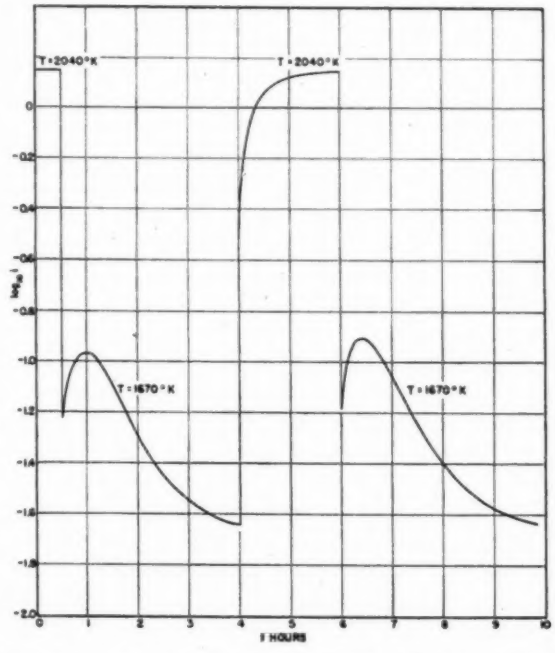

FIG. 5. Reversible change of activity with cathode temperature.

[4] D. A. Wright, Nature **160**, 129 (1947).
[5] T. E. Hanley, J. App. Phys. **19**, 583 (1948).
[6] M. O. Weinreich, Rev. Gen. Elec. **54**, 243 (1945).

tions on this effect will now be given in detail. If the cathode was kept at a temperature $T_1$ until its emission had reached a steady value then brought down to a lower value $T_2$, the emission tested intermittently at $T_2$ changed with time. If the temperature was again restored to $T_1$, the emission would build up gradually to its original equilibrium value. The process could then be repeated again. This reversible change is shown in Fig. 5 for $T_1=2040°K$ and $T_2=1670°K$. It must be emphasized that the anode voltage was applied only when a point was taken and only long enough for the measurement to be made. This is, therefore, an entirely different phenomenon from the decay of emission under applied voltage observed for oxide-coated cathodes. Figure 6 shows the good reproducibility of such curves taken at $T_2\sim1670°K$ for two values of $T_1$: 2320°K and 2040°K. Curves taken many days apart, during which time the cathode was subjected to different operating conditions, would check each other if only the cathode was initially brought to equilibrium conditions at the same $T_1$. This fact indicates that it was not the effect of residual gases in the tube. In a gettered tube gas effect cannot be expected to show such good reproducibility. Furthermore, it was found that such a curve could be interrupted for a long time, overnight, with the cathode cold; the emission would start again at the same value where the curve was interrupted and keep on changing in the same way as if there had been no interruption. Curves were taken at different $T_2$ values ranging from 1200°K to 1920°K

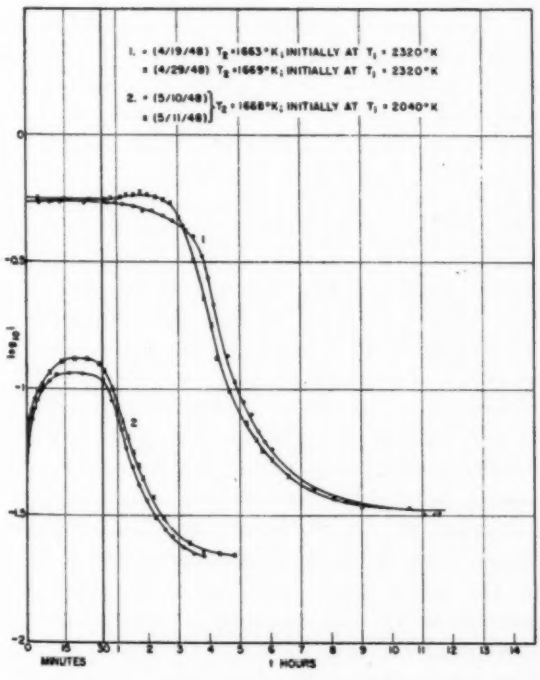


FIG. 6. Variation of emission with time after a decrease in cathode temperature.

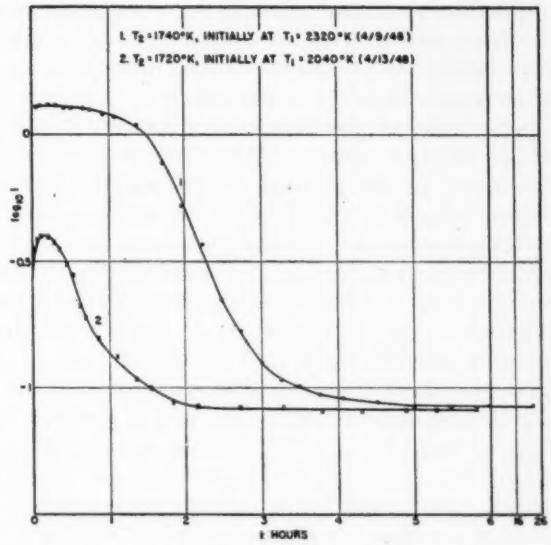


FIG. 7. Variation of emission with time after a decrease in cathode temperature.

for the same two values of $T_1$. Figures 7 and 8 give the curves for two other values of $T_2$. They show the same behavior. Comparing Figs. 6, 7 and 8 we see that with decreasing temperature the rate of change of emission decreases. All curves for $T_1=2040°K$ show initial rise, whereas for $T_1=2340°K$ they start off rather flat. Generally for the test temperature range 1200°K to 1920°K curves are similar in character. At very low testing temperatures, $T_2=1210°K$, the initial rise in curves for $T_1=2040°K$ could still be observed. The emission rose in 6 or 7 hours to a rather steady value, thereafter changed little in the course of 24 hours. The drop in emission observed at higher $T_2$ had either disappeared or proceeded at a very slow rate.

To follow the change in emission constants data for curves of $\log(i/T^2)$ *versus* $1/T$ were taken quickly at four points $A$, $B$, $C$ and $D$ of the run shown by curve 2, Fig. 8. These curves and the values of emission constants they yield are given in Fig. 9. They show that both $\varphi$ and $A$ underwent changes.

In a cathode with thoria and tungsten present together such reversible and reproducible changes of activity can naturally be ascribed to the production, diffusion and evaporation of free thorium. The effect of free thorium on the activity can be either a bulk effect or a surface effect. When dispersed inside thoria crystals free thorium may change $\varphi$ and $A$ of thoria as impurities in a semi-conductor. It may also change $\varphi$ and $A$ by being absorbed on the surface of the cathode. It may be argued that free thorium must be produced during the operation of the cathode, since if we have only a fixed amount of thorium as produced in the cathode originally during its fabrication, then without evaporation an

equilibrium condition should soon be established and there would then be no changes of activity. Evaporation should cause irreversible changes in the activity which are not observed.

The results of the following experiment, shown in Fig. 10, are significant for understanding the mechanism of the change of activity. After the emission reached a steady value at $T_1=2040°K$, the temperature of the cathode was lowered to $T_2=1670°K$ and changes of emission with time were observed. After an interval of time the cathode temperature was increased to $T_1$ for a very short time and brought back to $T_2$ again. In one of the curves the short time elevation to $T_1$ was made at the time when the emission at 1679°K had reached its peak. In the second curve this was made when the emission at 1670°K had passed beyond the peak. In both cases immediately after the short time at the higher temperature the emission was greatly reduced. It then partially recovered.

In the light of all the experiments we have arrived at the following picture of the mechanism of the cathode: (1) The variation of the activity is caused mainly by the variation of the amount of free thorium adsorbed on the surface. (2) Free thorium is produced through reduction of thoria by tungsten throughout the cathode. (3) The heat of reduction of thoria by tungsten is higher than the heat of evaporation of thorium from the cathode, which is in turn higher than the heat of diffusion of thorium in the cathode.

According to this hypothesis, when the temperature of the cathode is suddenly raised the amount of absorbed thorium will at first decrease on account of more rapid increase of evaporation than diffusion. However, the increased rate of thorium production more than compensates the increased rate of evaporation. Therefore, if the cathode is held at the higher temperature, eventually a higher thorium concentration will be built up, which in equilibrium supports a larger number of thorium atoms on the surface. The experimental facts: (1) Increasing the cathode temperature for a short time decreases the emission (Fig. 10). (2) After a decrease in cathode temperature there is an initial rise of emission with time (curves 2, Figs. 6, 7 and 8), and (3) Equilibrium emission activity is higher for higher temperature (compare initial emission values of curves 1 and 2, Figs. 6, 7 and 8), can thus be all explained. The fact that curves 1, Figs. 6, 7 and 8, do not show initial rise in emission, we assume to indicate that at $T_1=2320°K$ the amount of adsorbed thorium is close to a full monolayer.

Although in thoriated tungsten filaments considerable production of thorium takes place at temperatures higher than the temperatures reached by our cathode, the production should be much more rapid in our cathode than in a thoriated tungsten filament at the same temperature. Whereas a thoriated tungsten filament contains about 1 percent thoria our cathode is a mixture of 33 percent tungsten and 77 percent thoria.

The evaporation of thorium from tungsten is very slow at the temperatures at which changes of

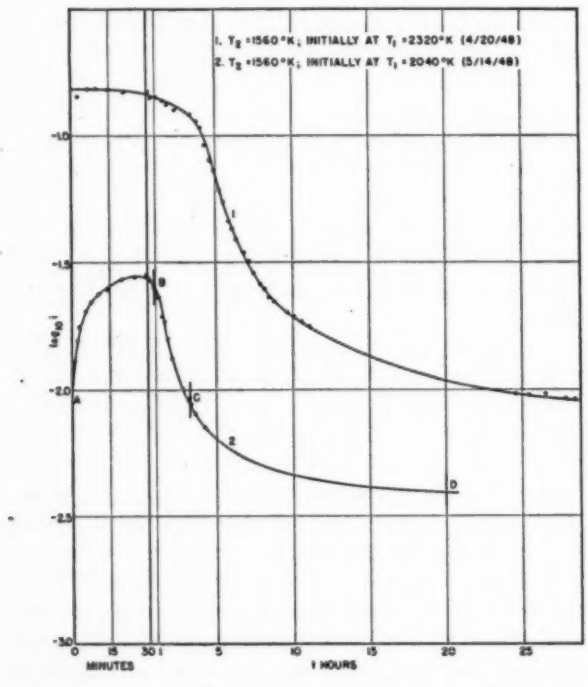


FIG. 8. Variation of emission with time after a decrease in cathode temperature.

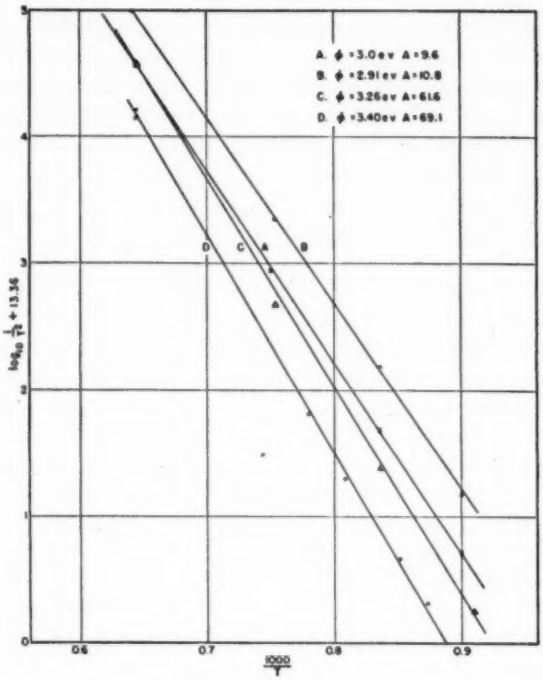


FIG. 9. Richardson plots for different states of the cathode.

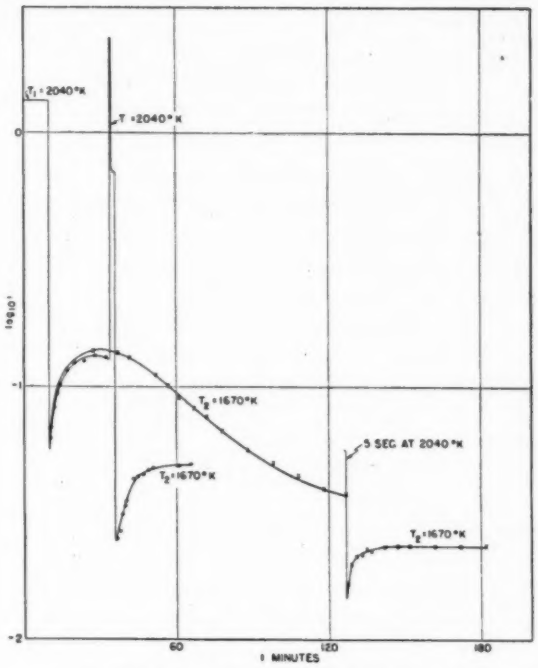


FIG. 10. Effect of short-time-elevation of the cathode temperature.

activity of our cathode were observed as shown by Figs. 6, 7, and 8. Also the heat of evaporation of thorium from tungsten is higher than the heat of reduction of thoria. But in our case thorium may be evaporated from surfaces of thoria rather than from tungsten. The data on thoriated tungsten do not necessarily contradict our hypothesis. However to substantiate our hypothesis an independent proof of evaporation of thorium during the change of cathode activity is desirable. The following experiment was made for this purpose.

## PROBE MEASUREMENT

The tube was rebuilt with the same cathode. The structure is the same as described previously except for the addition of a probe. Its arrangement is shown by the dotted lines in Fig. 2. It is a hairpin filament of 10 mil tungsten wire. Its temperature is calculated from the heating current and its emission measured at 1390°K, 1550°K and 1760°K checks with the calculated values. It serves to detect evaporation of materials from the cathode, which condensing on it changes its emission. Its emission could always be brought back to normal values by flashing it at 2460°K for a sufficient period of time. A few seconds of flashing was enough to bring the emission almost back to normal values. It was found necessary sometimes to keep the probe at the flashing temperature for many minutes before its emission at the low testing temperature, 1390°K, was quite restored to its normal value. This result is probably due to the fact that the top and the base of the hairpin are not far away enough from the cathode. The small amounts of material deposited on the tips of the lead wires, the spring, and the short lengths of filament near the contacts with the wires and the spring were not easily driven off and kept to diffuse on to the filament. In testing the probe emission, the cathode, the anode and the guard rings were tied together to serve as the anode. The emission was measured with a galvanometer the sensitivity of which is $4.4 \cdot 10^{-10}$ amp./mm.

## EVAPORATION OF THORIUM

In the following experiments the probe was first cleaned by being kept at 2460°K for several minutes (with cathode at room temperature) until its emission at the testing temperature $T = 1250$°K gave the normal value of the clean surface. It was found that after glowing the cathode for six hours at 1770°K the probe emission at the testing temperature would not change appreciably, if the cathode had been previously brought to equilibrium at 1770°K. However, if the cathode had been previously heated at 2220°K, glowing the cathode for two hours at 1770°K increased the probe emission 20,000 times. These experiments prove conclusively that the increase of probe emission in the second case must be due to free thorium evaporated from the cathode. The only other material that could come from the cathode to enhance the probe emission is thoria itself. But the evaporation of thoria depends upon the cathode temperature only and should not show such decided dependence on the previous treatment of the cathode. The conclusion is that considerable reduction of thoria takes place in the cathode at 2220°K building up a concentration of free thorium which keeps on evaporating at 1770°K. This is then a direct evidence of the correctness of our picture of the mechanism in the cathode.

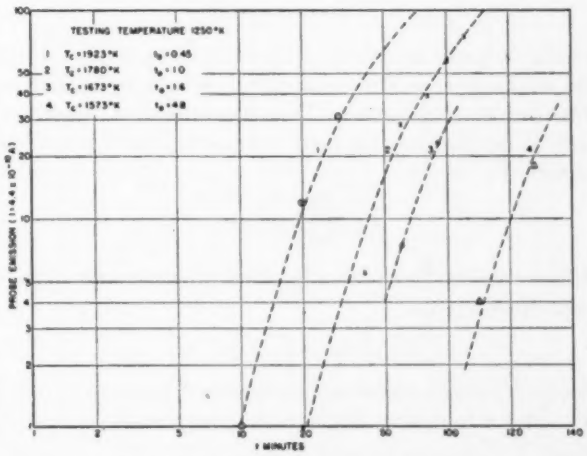
TESTING TEMPERATURE 1250°K
1 $T_0$ = 1923°K $t_0$ = 0.45
2 $T_0$ = 1780°K $t_0$ = 1.0
3 $T_0$ = 1673°K $t_0$ = 1.6
4 $T_0$ = 1573°K $t_0$ = 4.8
PROBE EMISSION (1 = 4.4 × 10⁻¹⁰ A)
t MINUTES

FIG. 11. Probe emission *versus* time of glowing the cathode.

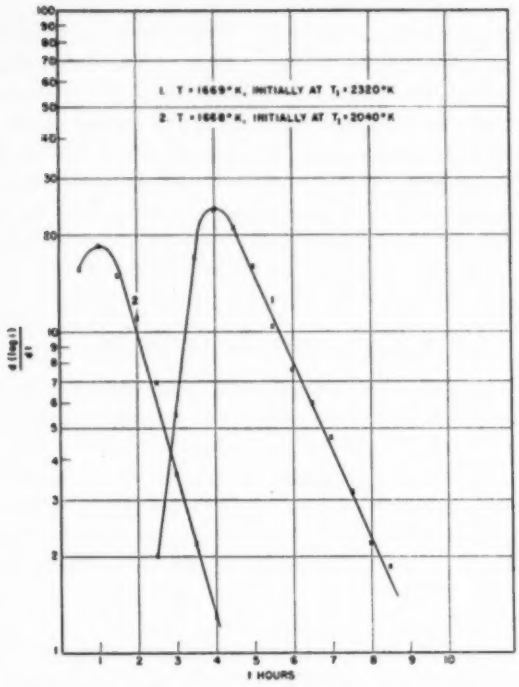


FIG. 12. $[d(\log i)/dt]$ *versus* time, after the cathode temperature is brought down from a high value.

Figure 11 shows the probe emission at 1250°K as function of the time interval of glowing the cathode. Points taken for four different glowing temperatures of the cathode can be fitted by a single curve which is shifted horizontally a suitable distance for each cathode temperature. From these curves we get relative time intervals $t_0$ necessary to glow the cathode at the different temperatures to obtain a certain value of the probe emission. Thus $1/t_0$ is the relative rate of thorium evaporation.

The rate of thorium evaporation can also be estimated from other experiments. Assuming the rate of evaporation to be proportional to the number of atoms on the surface we have

$$dN/dt = DG - EN \tag{2}$$

where $N$ is the number of thorium atoms per cm² of the surface, $G$ is the concentration gradient of free thorium at the surface, $D$ is the diffusion constant, and $E$ is the constant for evaporation. It can be assumed:

$$D = D_0 e^{-(E_d/kT)}, \qquad E = E_0 e^{-(E_e/kT)}, \tag{3}$$

In the case of thorium adsorbed on tungsten, it has been shown that[7]

$$\log(i/i_0) = L(N/N_c), \tag{4}$$

where $N_c$ is the number of adsorbed thorium atoms per cm² when the surface is completely covered,

[7] W. H. Brattain and J. A. Becker, Phys. Rev. **43**, 428 (1933).

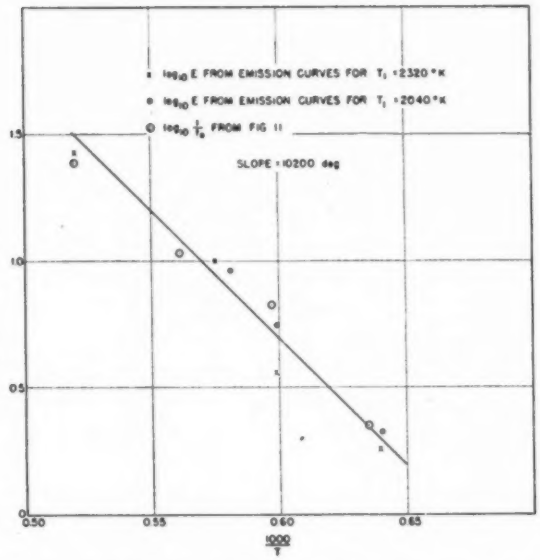
$\log_{10} E$ FROM EMISSION CURVES FOR $T_1$ = 2320°K
$\log_{10} E$ FROM EMISSION CURVES FOR $T_1$ = 2040°K
$\log_{10} \frac{1}{t_0}$ FROM FIG 11
SLOPE = 10200 deg
$\frac{1000}{T}$

FIG. 13. Determination of the heat of evaporation of free thorium.

$i$ and $i_c$ are the emission currents corresponding to $N$ and $N_c$ respectively. The factor $L$ depends on $N$ but may be taken as constant for limited ranges of $N/N_c$. By analogy we shall assume

$$\log i \propto (N + \text{const.}). \tag{5}$$

When the cathode is brought down from a higher temperature to a temperature at which the reduction of thoria is negligible, the concentration of free thorium becomes depleted with time. Eventually $G$ becomes small and the evaporation term in Eq. (2) becomes predominant over the diffusion term:

$$dN/dt \approx -EN. \tag{6}$$

In view of Eq. (5) we have

$$\frac{d(\log i)}{dt} = \frac{dN}{dt} \propto e^{-Et}, \tag{7}$$

$$\log \frac{d(\log i)}{dt} = \text{const.} - Et. \tag{8}$$

Figure 12 shows curves of $[d(\log i)/dt]$ *versus* $t$ in semilogarithmic plot. They are calculated from the experimental curves of Fig. 7. When the experimental curves begin to fall quickly (according to our interpretation, thorium concentration has been depleted enough for the evaporation to predominate) these plots do become straight lines. According to our theory the slope of such straight line being equal to $E$ should be determined by $T_2$ independent of $T_1$.

Figure 13 gives $\log E$ and $\log 1/t_0$ (from Fig. 11) plotted against $1/T$. The two sets of points $\log E$ and $\log 1/t$ can be shifted vertically with respect to

each other. All the points should fall on a straight line according to Eq. (3). The straight line drawn to fit these points is given by

$$\log_{10}E = \text{const.} + (10200/T). \qquad (10)$$

The slope of this line multiplied by 4.575 gives the heat of evaporation to be 46600 calories per g-atom. Although this slope is not determined very accurately due to the scattering of the points, it serves to indicate the order of magnitude. In view of the fact that the points were determined from experiments made many days apart and the probe measurements were made with a different tube, the scattering may be considered not too bad. The fact that $\log 1/t_0$ and $\log E$ derived from different kinds of experiments can be made to fall around the same line lends support to our theory. That the heat of evaporation obtained is much lower than the value for evaporation of thorium from tungsten ($4.575 \times 44500$ cal./g-atom)[8] indicates that thorium is evaporated from the surface of thoria.

## EVAPORATION OF THORIA

At cathode temperatures above 1900°K the following phenomena were observed in the probe emission. After the cathode had been glowed for a certain period of time the probe emission measured at 1390°K was above its normal value. When the probe temperature was raised to 1550°K the emission decayed fairly quickly and might even drop below the normal emission. But flashing for a few seconds at 1970°K made the emission tested at 1390°K rise again. With successive flashing the emission went through a peak then fell approaching its normal value. This is shown in Fig. 14. The measurements were taken after the cathode was glowed at 2130°K for 5 minutes with the probe unheated.

The fall of emission at 1550°K is apparently due to rearrangement by diffusion of materials deposited on the probe. The rise of the emission during the flashing of the probe we assume to be due to reduction of thoria deposited on the probe from the cathode. Figure 15 shows the probe emission, measured at 1390°K while the probe was flashed at 1970°K, after the cathodes had been glowed for different lengths of time at 2030°K. The bottom line corresponds to about ten times the normal emission. The time required for the emission to fall to this value is seen to be approximately proportional to the time for which the cathode had been glowed. The same result was observed at 2230°K, 2130°K and 1940°K. We take this time as a measure of the amount of thoria deposited on the probe. Below the value of about ten times the normal emission the curves became irregular. Materials deposited on the ends of the filament may be the cause of this irregularity. With this measure of the amount of deposited thoria we can compute from data taken for different temperatures $T$ of the cathode (curves similar to those shown in Fig. 15) the time, $t_e$, required for glowing the cathode to deposit a certain amount of thoria on the probe. Figure 16 shows the straight line relation between

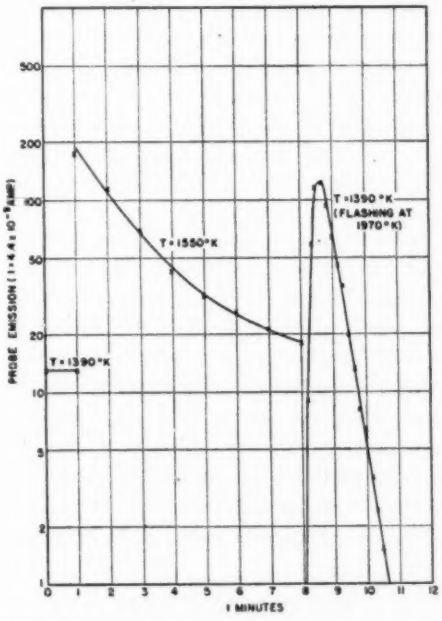


FIG. 14. Variation of probe emission after a period of heating the cathode. Emission measured at 1390°K, 1550°K, and at 1390°K while being kept at 1970°K.

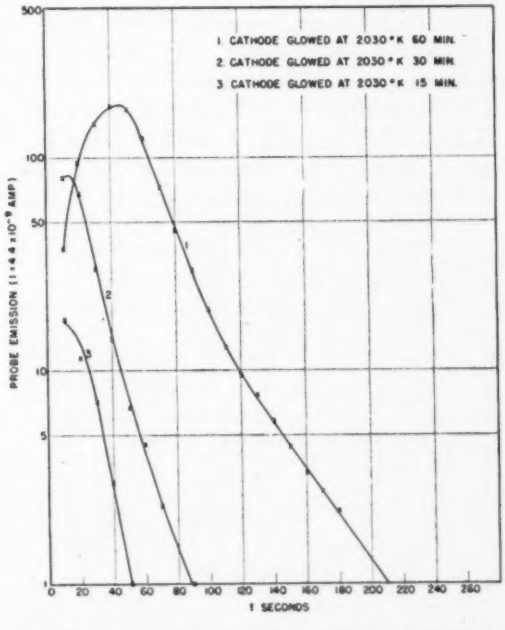


FIG. 15. Time variation of probe emission at 1970°K (testing temperature 1390°K) after glowing the cathode at 2030°K for different intervals of time.

[8] I. Langmuir, Phys. Rev. **22**, 357 (1923).

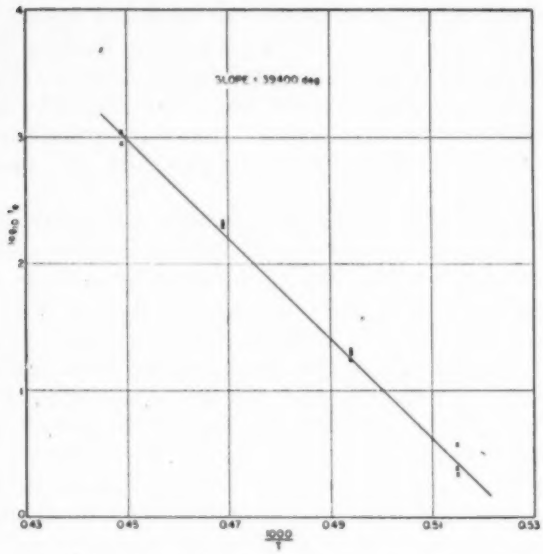


FIG. 16. Determination of the heat of evaporation of thoria.

$\log_{10} i_e$ and $1/T$. The equation of this line is:

$$\log_{10} i_e = \text{const.} - (39{,}400/T). \qquad (11)$$

The heat of evaporation of thoria is thus 184,000 calories per g-molecule.

## DIFFUSION OF THORIUM

Equations (2) and (5) give:

$$[d(\log i)/dt] \propto DG - EN.$$

By allowing the cathode to come to equilibrium at a fixed temperature $T_1$ we established the same values of $G$ and $N$. The initial values, $|d(\log i)/dt|_{t=0}$, when the cathode is brought down from $T_1$ to different temperatures should then be a combination of two exponential functions of $1/T$ determined by $D$ and $E$ according to (3). Figure 17 gives the curve $\log|d(\log i)/dt|_{t=0}$ *versus* $1/T$ for $T_1 = 2040°\text{K}$. These data are taken from curves similar to curves 2 of Figs. 6–8. For temperatures below 1700°K the curve is straight down to 1100°K. At the lower temperature the evaporation will be very slow. The slope of this line corresponds then to the heat of diffusion. The equation of this line is

$$\log_{10}|d(\log i)/dt|_{t=0} = \text{const.} - (6840/T)$$

giving the heat of diffusion to be 31,900 calories per g-atom. At 1700°K the curve begins to bend due to the increased predominance of evaporation.

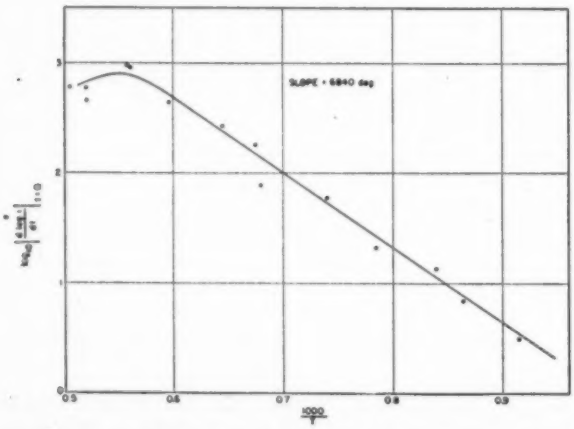


FIG. 17. Determination of the heat of diffusion of free thorium.

## CONCLUDING REMARKS

The addition of tungsten powder for the purpose of simplifying the heating problem seems to retain all advantages of the sintered thoria cathode, except for a somewhat higher work function. Thus at 2040°K the maximum space charge limited emission at the bend of the curve in Fig. 3 is 0.9 amp./cm², whereas Hanley's data for thoria coated on tungsten give about 4 amp./cm². This is due to the fact that part of the surface is occupied by tungsten which has a high work function.

Our theory seems to give satisfactory explanations to the observed changes of activity with temperature. The probe measurements support our theory in confirming the evaporation of thorium. The estimated values of the heat of evaporation, 46,000 cal./g-atom, is higher than the estimated value of heat of diffusion, 31,900 cal./g-atom, and is lower than the known value of the heat of reduction of thoria, 138,000 cal./g-atom.[8] This is in accordance with the assumption of our theory.

It appears from our results that all thoria cathodes, where hot tungsten is in contact with thoria, should show changes of activity with temperature variation of the cathode. Indeed, Weinreich[6] observed that the emission of thoria coated on tungsten at 1800°K starting at a high value decayed with time, if the cathode had been previously flashed at a temperature above 2600°K. Thus in his case change of activity does occur, but only by flashing the cathode at temperatures above 2600°K. It is significant, however, that, similar to our case, the emission right after flashing was much higher than the equilibrium value at the lower temperature, 1800°K. A thoriated tungsten filament would be completely deactivated at the flashing temperature (above 2600°K) and would begin to activate at the lower temperature. As discussed above, this is, according to our theory, because the rate of thorium evaporation from thoria rises with temperature more slowly as compared with its evaporation from tungsten.

## ACKNOWLEDGMENTS

The author wishes to thank Professor W. B. Nottingham for many discussions. He also wishes to acknowledge his indebtedness to Professor S. T. Martin for his help in carrying out this work.

# High Temperature Furnace for Electron Diffraction Studies of Thin Films

EILEEN I. ALESSANDRINI
*Research Laboratory, General Electric Company, Schenectady, New York*
(Received January 18, 1949)

A furnace has been designed which enables one to carry out electron diffraction studies of thin films at high temperatures. Electron diffraction patterns are obtainable at temperatures up to 1100°C. It is possible to bring the furnace up to 1500°C, but at temperatures over 1100°C the photographic plates are fogged by light emitted from the furnace. With the use of this furnace, recrystallizations and phase changes in thin metal and alloy films can be studied at the temperatures where these phenomena occur.

## GENERAL DESCRIPTION

A HIGH temperature furnace has been designed for use in making electron diffraction studies of thin metallic films of metals and alloys. It has been used with a laboratory model, but can be easily adapted to any electron diffraction instrument. The furnace is constructed so that it is possible to obtain transmission patterns of these films at temperatures up to 1100°C. The furnace can be mounted in place of the sample holder in an electron diffraction instrument, and vertical and horizontal motions of the thin film mount are provided, so that any chosen portion of the sample can be brought into the electron beam.

The furnace mounting consists of a vacuum seal through which it is possible to move the specimen, heat it and observe temperatures by means of a thermocouple. The films are heated by a helically wound tungsten wire, which surrounds the outer edge of the circular film holder. A direct current is applied to the tungsten wire, and the heat is radiated to the film. A bank of four storage batteries is sufficient to obtain temperatures up to 1100°C. In order to obtain still higher temperatures, electron bombardment of the specimen can be accomplished. By connecting the positive terminal of one or more 90-volt batteries to the specimen holder and the negative terminal to one leg of the tungsten wire, the specimen can be bombarded with electrons and heated to temperatures over 1500°C. A radiation shield of tantalum is used to keep heat losses down to a minimum. Any amount of gases evolved when the furnace is run at high temperatures is readily taken care of by the pumps of the diffraction instrument.

Specimens for use in the high temperature furnace are prepared by evaporation. A metal or metals can be evaporated onto a plastic film. This film of plastic plus metal, in turn, is supported by a fine meshed screen which is placed in the furnace. Before films are mounted in the furnace for high temperature studies, the plastic material is removed by a solvent, and therefore no contamination of the metal film occurs.

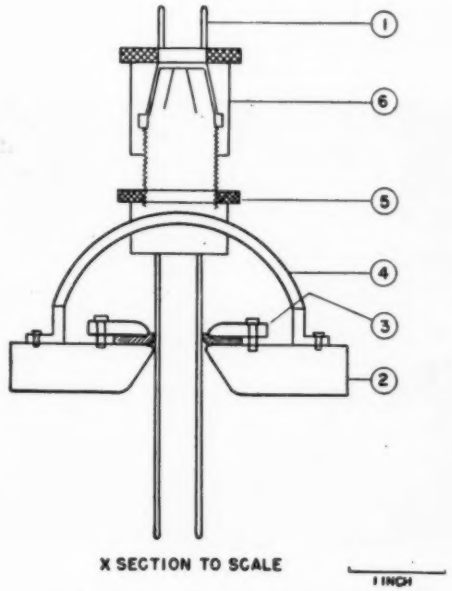


FIG. 1. Furnace manipulator.

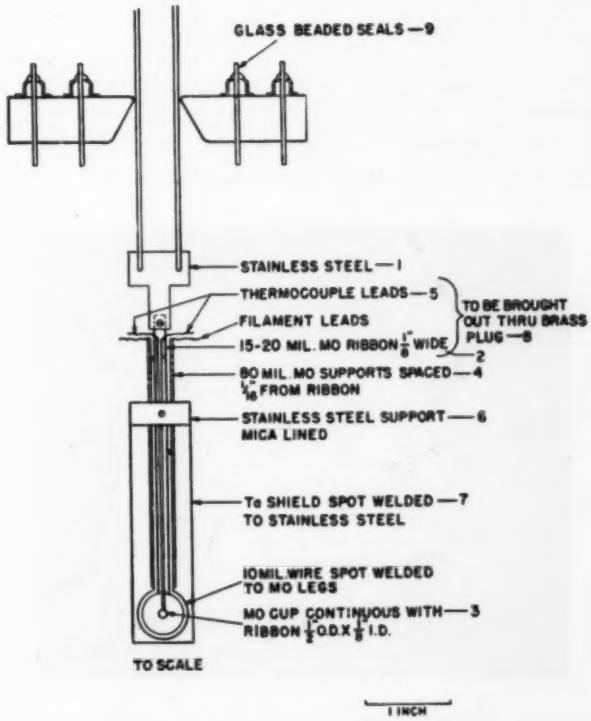
GLASS BEADED SEALS—9
STAINLESS STEEL—1
THERMOCOUPLE LEADS—5
FILAMENT LEADS
15-20 MIL MO RIBBON 1/8" WIDE
TO BE BROUGHT OUT THRU BRASS PLUG—8
2
80 MIL MO SUPPORTS SPACED—4 1/16" FROM RIBBON
STAINLESS STEEL SUPPORT—6 MICA LINED
To SHIELD SPOT WELDED—7 TO STAINLESS STEEL
10MIL WIRE SPOT WELDED TO MO LEGS
MO CUP CONTINUOUS WITH—3 RIBBON 1/2 O.D. X 1/4 I.D.
TO SCALE
1 INCH

FIG. 2. High temperature furnace.

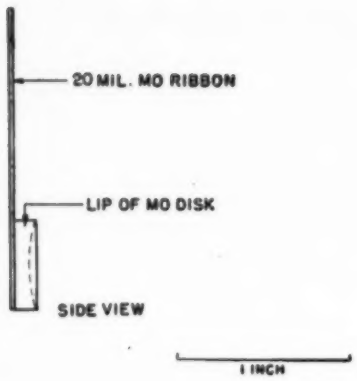


FIG. 3. Detail of furnace disk.

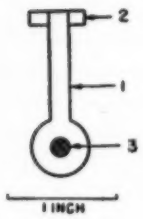


FIG. 4. Specimen holder.

## SPECIFICATIONS

The furnace mounting is arranged to pass through a vacuum seal, making it possible to move the specimen from outside. This arrangement is illustrated in Fig. 1. A stainless steel tube (1) has been made to fit through a heavy brass disk (2). The vacuum is maintained by means of a Wilson seal (3). The steel tube is used to give support to the specimen holder and also to make it possible to move the specimen vertically and horizontally in the electron beam. Additional support is given the tube by a $\frac{1}{8}$-inch track (4). A slide and knurled nut (5) are used to secure horizontal positioning of the specimen, and a split chuck and knurled nut (6) are used to hold the specimen in any desired vertical position. The two filament terminals and thermocouple leads are brought out through the heavy brass disk (not shown in drawing). Glass beaded seals are used to insulate them.

On the end of the stainless steel tube which extends into the vacuum (Fig. 2), a stainless steel plug (1) is copper brazed. A slit and screw are provided which hold a 20 mil, $\frac{1}{8}$-inch wide molybdenum lath (2). This ribbon lath terminates in a disk $\frac{1}{2}$ inch in diameter with a $\frac{1}{8}$-inch aperture in its center. This disk has a lip which is $\frac{3}{32}$ inch in width, as shown in Fig. 3.

A 10 mil helically wound tungesten wire (3) encircles the lip of this disk without touching it. This wire serves as the heating element of the furnace. This heating element is supported by two 80 mil molybdenum wires (4). Copper braided wires extend from these supports to the terminals (8). The thermocouple wires (5), of 10 mil Chromel and Alumel, are spot welded to the molybdenum disk at the edge of the aperture. The terminals (9) for these wires extend to the outside of the furnace through glass beaded seals. All of these leads are held rigidly in place along the molybdenum lath by means of a mica lined stainless steel block (6), which consists of two pieces screwed down tightly against the lath.

In order to mount specimens in this furnace, a molybdenum disk $\frac{15}{32}$ inch in diameter and having an aperture $\frac{1}{8}$ inch in diameter was made (Fig. 4). This disk is 20 mils in thickness and has a $\frac{1}{8}$ inch ribbon extending $1\frac{1}{2}$ inches from an edge (Fig. 4 (1)). At the edge of the ribbon a piece of 5 mil molybdenum was spot welded to form tabs (Fig. 4 (2)) for securing the disk to the furnace. This specimen holder fits flush against the molybdenum disk of the furnace, and the two $\frac{1}{8}$-inch apertures are concentric. Across the aperture of the specimen holder a 150-mesh platinum-rhodium screen (Fig. 4 (3)) was spot welded. This screen serves as the film support.

A radiation shield of tantalum is used to cut down heat loss to the instrument. This shield is a piece of tantalum which is spot welded to form a cylinder (Fig. 2(7)) $\frac{3}{4}$ inch in diameter. A disk of tantalum is spot

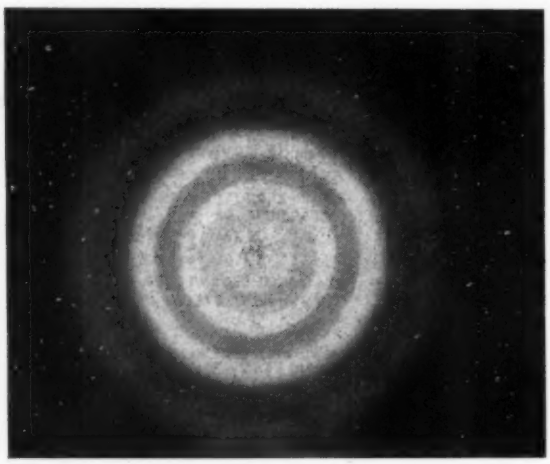

FIG. 5. Germanium evaporated on a high temperature furnace mesh.

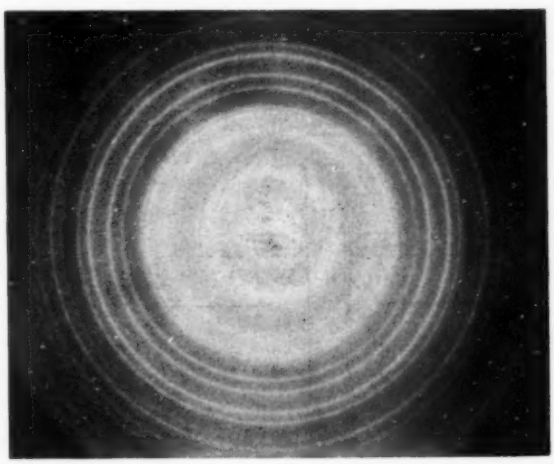

FIG. 6. Germanium heated to 350°C showing recrystallization.

welded across the bottom and three tabs were cut out at the top. This shield fits around the furnace and is secured by means of the tabs to the stainless steel, mica lined, supporting block (Fig. 2(6)). Two $\frac{3}{16}$-inch apertures were made for the entrance and exit of the electron beam.

A film for diffraction studies is mounted in the furnace, the filament of which is heated with the use of storage batteries. Twelve cells are connected in series, and the amount of current through the filament is controlled by a variable resistor. In this way, specimens can be brought to temperature gradually. For temperatures over 1000°C, the specimens are bombarded with electrons. Dry batteries are connected across the specimen holder and one filament terminal. If the specimen is made positive, electrons from the filament bombard the molybdenum disk and cause an increase in temperature.

## RESULTS

Several preliminary experiments have been carried out using this high temperature furnace. Germanium was evaporated on Formvar. The Formvar was removed by placing the film in a solution of ethylene dichloride. The metal film was then mounted in the high temperature furnace and a diffraction pattern was taken (Fig. 5). This pattern shows four diffraction rings. The germanium was heated to 350°C, and recrystallization occurred, as shown by the increased number and sharpness of the rings (Fig. 6).

## ACKNOWLEDGMENT

The author wishes to thank Professor Charles Bachmen, Physics Department, Syracuse University for his many suggestions regarding this work.

# A Method of Equivalent Linearization for Non-Linear Oscillatory Systems with Large Non-Linearity

CARL A. LUDEKE
*University of Cincinnati, Cincinnati, Ohio*
(Received October 15, 1948)

This paper introduces a method, based partially on existing non-linear theory and partially on experimental results, for transforming a non-linear oscillatory system into an equivalent linear one, regardless of the amount of non-linearity in the original system.

## INTRODUCTION

THE applied physicist, concerned with non-linear forced damped oscillations, is interested in the wave form, the frequency-amplitude relationship, and the dependence of the maximum amplitude on damping. In most practical problems, the damping force and the applied force are so small compared to the restoring force that if questions concerning these three points of interest can be satisfactorily answered, the complete behavior of the system is essentially established. Although methods are available for answering such questions if the non-linearity is small, this author believes that more should be said about the application of these methods to systems with large non-linearity.

Of the methods available two have been chosen as characteristic of all others, and the difficulties involved in applying them to systems with large non-linearity are cited. In an attempt to remove some of these difficulties, the author uses experimental results obtained by means of an apparatus previously reported.[1]

## THEORY

Before examining these two methods, let us consider the differential equation,

$$I\ddot{\theta}+k_1\theta+k_3\theta^3=0, \tag{1}$$

and assume it has a solution

$$\theta=A\sin\omega t+pA\sin3\omega t, \tag{2}$$

in which $p$ is so small that terms of the order of $p^2$ may be neglected. Under what conditions is this assumption approximately correct?

Upon substituting Eq. (2) into Eq. (1) and neglecting terms of the order of $p^2$ we have

$$[(k_1-I\omega^2)A+3k_3A^3/4]\sin\omega t +[(k_1-9I\omega^2)pA+3pk_3A^3/2-k_3A^3/4]\sin3\omega t -(3pk_3A^3/2)\cos2\omega t\sin3\omega t=0. \tag{3}$$

Obviously Eq. (3) can be true for all values of $t$ only if $p=0$, in which case $k_3=0$, and $\omega^2=k_1/I$; this is the well known linear case. However, with $p\neq0$, we can still find an expression for $\omega^2$ which will satisfy Eq. (3) for a particular value of $t$. When restricted to one value of $t$, the usual choice is $t=2\pi/\omega$, in which case we have the relationship,

$$I\omega^2=[(1-p)/(1-9p)]k_1 +[(1-3p)/(1-9p)(1-p^2)]k_3\theta_m^2, \tag{4}$$

in which $\theta_m$ is the maximum value of $\theta$.

If the solution represented by Eq. (2) exists with small values of $p$ regardless of the magnitude of the terms in Eq. (1), then it is possible to write Eq. (4) in the approximate form,

$$I\omega^2\simeq k_1+k_3\theta_m^2\simeq k_1+k_3A^2. \tag{5}$$

If, however, a small value of $p$ demands a small value of $k_3$, then Eq. (3) reduces to

$$[-A\omega^2I+k_1A+3k_3A^3/4]\sin\omega t=0, \tag{6}$$

which can be satisfied for all values of $t$, by the expression

$$I\omega^2=k_1+3k_3A^2/4. \tag{7}$$

A consideration of Eqs. (5) and (7) indicates (a) that the form of the expression relating frequency and amplitude might remain the same for all values of non-linearity, and (b) if the solution remains essentially sinusoidal for large non-linearities, then the coefficient of the term $k_3A^2$ is 3/4 when the non-linearity is small and approximately 1 when the non-linearity is large.

If the right side of Eq. (1) is replaced by $P_0\sin\omega t$, Eq. (1) becomes the so-called Duffing[2] equation, and Eqs. (5) and (7) are replaced by

$$I\omega^2\simeq k_1+k_3\theta_m^2-P_0/\theta_m\simeq k_1+k_3A^2-P_0/A, \tag{8}$$

and

$$I\omega^2=k_1+3k_3A^2/4-P_0/A \tag{9}$$

respectively.

In order to determine whether or not the solutions are approximately sinusoidal when the non-linearity is large, and in order to determine the best coefficient for the term $k_3A^2$ let us consider a well-known method[3] available for solving such problems when the non-linearity is small. In reference 3, p. 135, this

[1] C. A. Ludeke, J. App. Phys. **20**, 600 (1949).

[2] G. Duffing, "Sammlung Vieweg," Heft 41/42, 1918.

[3] N. Minorsky, *Introduction to Non-Linear Mechanics* (J. W. Edwards, Ann Arbor, Michigan, 1947).

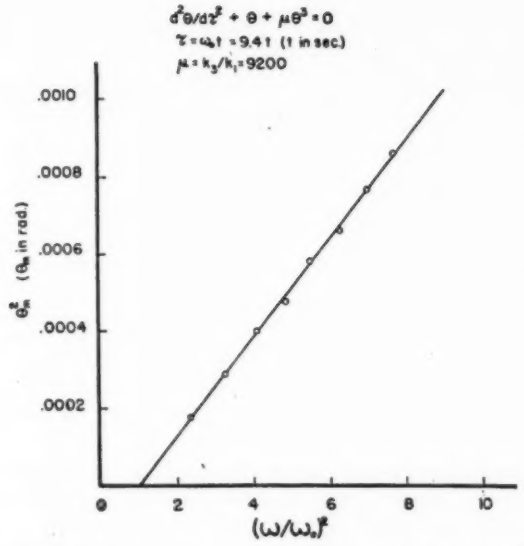


FIG. 1. Frequency-amplitude data for a typical non-linear free oscillation.

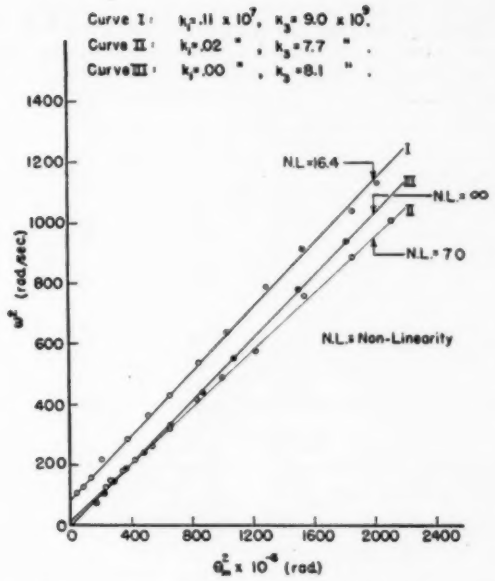


FIG. 2. Frequency-amplitude data for several typical non-linear forces oscillations.

restriction of small non-linearity is explicitly stated. In article 71, p. 222, Eq. (1) is considered, and from Eqs. (71.5) and (71.6) on p. 223 it is evident that these equations which give the solution and the frequency-amplitude relationship are only true when the non-linearity is small. Thus the applied physicist would not find the method discussed in reference 3 suitable for problems with large non-linearity even though it is a method capable of yielding approximations of higher orders.

As a specific example, consider Eq. (1) written in the form

$$d^2\theta/d\tau^2+\theta+\mu\theta^3=0, \qquad (10)$$

in which $\tau=(k_1/I)^{\frac{1}{2}}t=\omega_0 t$ and $\mu=k_3/I\omega_0^2=k_3/k_1$ $=9200$. With this large value of $\mu$, the higher power terms in Eq. (71.6) of reference 3 would predominate and hence this equation could not be used. The actual experimental results are shown in Fig. 1, from which we see that it is only necessary to retain the term in the square of the amplitude. Thus, as clearly stated in reference 3 this method should not be used for systems with large non-linearity.

TABLE I. Experimentally determined values of $C$, the coefficient of the term $\beta\theta_m^2$, for systems having a restoring torque $k_1\theta+k_3\theta^3$.

| Spring identification | $C$ | $k_1$ (g cm/rad.) | $k_3$ (g cm/rad.$^3$) |
|---|---|---|---|
| 13;0–0 | 1.01 | $1.27\times10^7$ | $15.2\times10^9$ |
| 13;0–1 | 1.01 | 1.17 | 14.0 |
| 13;0–2 | 0.93 | 1.14 | 14.0 |
| 13;0–3 | 0.94 | 1.10 | 11.9 |
| 13;0–4 | 0.89 | 1.06 | 11.9 |
| 13;0–5 | 0.84 | 1.03 | 11.1 |
| 13;0–6 | 0.99 | 0.98 | 9.8 |
| 13;0–7 | 0.95 | 0.94 | 10.3 |
| 13;0–8 | 0.95 | 0.92 | 9.7 |
| 20;0–3 | 0.88 | 0.11 | 9.0 |
| 20;0–4 | 0.90 | 0.02 | 7.7 |
| 20;0–2 | 0.94 | 0.00 | 8.1 |
| 20;0–8*A* | 0.83 | 0.04 | 6.9 |
| 20;0–1 | 0.93 | 0.07 | 9.3 |
| 20;0–8*B* | 0.87 | 0.49 | 7.0 |
| 20;0–0 | 0.86 | 0.10 | 10.2 |
| 20;0–2*A* | 0.82 | 0.15 | 9.7 |
| 20;0–2*B* | 0.86 | 0.68 | 8.6 |
| 21;0–0 | 0.87 | 1.01 | 16.3 |

Consider now the work of Friedrichs and Stoker,[4] p. 36, in which we encounter an iteration method which does not depend explicitly upon the magnitude of the non-linearity but which does start with the free linear oscillation as the initial solution. Indication that the magnitude of the non-linearity is involved is given by the author of reference 4 on p. 39 where he states that the frequency-amplitude relationship can be expected to be accurate for not too large values of the term containing the square of the amplitude. Since the method under discussion can be applied repeatedly, let us consider the results of two such applications on the Duffing equation when it is written in the form,

$$\ddot{\theta}+\omega_0^2\theta+\beta\theta^3=F\cos\omega t, \qquad (11)$$

in which $\omega_0^2=k_1/I$, $\beta=k_3/I$ and $F=P_0/I$.

The first application yields as a solution for Eq. (11) the expression

$$\theta_1=A\cos\omega t + [\beta A^3/32(\omega_0^2+3\beta A^2/4-F/A)]\cos3\omega t. \qquad (12)$$

The first approximation to the frequency-amplitude relationship is

$$\omega_1^2=\omega_0^2+3\beta A^2/4-F/A. \qquad (13)$$

[4] K. O. Friedrichs and J. J. Stoker, "Introduction to Non-Linear Mechanics," New York University Lecture Notes.

TABLE II. Results of the harmonic analysis of three typical wave forms. (Note that the third harmonic is the only higher harmonic present to any appreciable amount, and that its amplitude decreases as the non-linearity decreases.)*

| Spring constants | $n$ | $a_n$ | $b_n$ | Non-linearity | Equation (approximately) |
|---|---|---|---|---|---|
| | | $\theta = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega t + \sum_{n=1}^{\infty} b_n \sin n\omega t$ | | | |
| | 1 | 0.00 | −2.34 | 8.67 | |
| $k_1 = 0.13 \times 10^7$ | 2 | 0.00 | 0.02 | | $\theta = -2.34 \sin\omega t$ |
| $k_3 = 11.95 \times 10^9$ | 3 | 0.00 | 0.12 | | $+0.12 \sin 3\omega t$ |
| | 4 | −0.01 | 0.00 | | |
| | 5 | 0.01 | 0.00 | | |
| | 6 | 0.01 | 0.01 | | |
| | 7 | 0.02 | 0.01 | | |
| | 8 | 0.01 | 0.01 | | |
| | 9 | 0.01 | 0.01 | | |
| | 10 | 0.01 | 0.01 | | |
| | 1 | −0.11 | 2.79 | 1.97 | |
| $k_1 = 0.60 \times 10^7$ | 2 | −0.05 | 0.00 | | $\theta = 2.79 \sin\omega t$ |
| $k_3 = 11.95 \times 10^9$ | 3 | 0.01 | −0.07 | | $-0.07 \sin 3\omega t$ |
| | 4 | 0.01 | 0.01 | | |
| | 5 | 0.00 | 0.02 | | |
| | 6 | 0.01 | 0.01 | | |
| | 7 | 0.03 | −0.01 | | |
| | 8 | 0.00 | 0.01 | | |
| | 9 | 0.00 | 0.00 | | |
| | 10 | 0.00 | 0.01 | | |
| | 1 | 0.06 | 3.98 | 0.98 | |
| $k_1 = 1.48 \times 10^7$ | 2 | −0.07 | −0.04 | | $\theta = 3.98 \sin\omega t$ |
| $k_3 = 11.95 \times 10^9$ | 3 | −0.03 | −0.07 | | |
| | 4 | −0.01 | 0.03 | | |
| | 5 | −0.01 | 0.01 | | |
| | 6 | 0.00 | 0.00 | | |
| | 7 | −0.02 | −0.02 | | |
| | 8 | 0.00 | 0.00 | | |
| | 9 | 0.00 | −0.01 | | |
| | 10 | 0.01 | 0.01 | | |

* Note: (1) $a_n$ and $b_n$ are measured in inches with the wave-length equal to 6 inches. (2) The smallest reading on the analyzer was 0.01 inch. (3) From theoretical considerations there should be no even harmonics.

A second application of the process yields as a solution

$$\theta_2 = A\cos\omega t - (1/8\omega^2)\{[(\omega^2 w_0^2)/32\omega_1^2 - 1/4]\beta A^3 - 3\beta^2 A^5/64\omega_1^2 - 3\beta A^9/131072\}\cos 3\omega t + (1/8\omega^2)[\beta^2 A^5/128\omega_1^2 + \beta^3 A^7/4096\omega_1^4]\cos 5\omega t + (1/16\omega^2)[\beta^3 A^7/4096\omega_1^4]\cos 7\omega t - (1/320\omega^2)[\beta^4 A^9/32768\omega_1^6]\cos 9\omega t \quad (14)$$

while the frequency-amplitude equation becomes,

$$\omega_2^2 = \omega_0^2 + (3\beta A^2/4)[1 + \beta A^2/32\omega_1^4] - F/A. \quad (15)$$

Let us examine Eqs. (12), (13), (14), and (15) when the non-linearity is so large that Eq. (13) can be written

$$\omega_1^2 \simeq 3\beta A^2/4. \quad (16)$$

Equation (12) then becomes

$$\theta_1 \simeq A\cos\omega t + (A/24)\cos 3\omega t, \quad (17)$$

which consists predominantly of the term $A\cos\omega t$, and which does not contain $\beta$ explicitly. Similarly, Eqs. (14) and (15) become,

$$\theta_2 \simeq A\cos\omega t + (3/64)A\cos 3\omega t + (1/600)A\cos 5\omega t + (1/25248)A\cos 7\omega t + (4.43\times 10^{-6})A\cos 9\omega t, \quad (18)$$

and

$$\omega_2^2 = \omega_0^2 + (3/4)(397/384)\beta A^2 - F/A \simeq (3/4)(397/384)\beta A^2. \quad (19)$$

From Eqs. (18) and (19) we see that even for large non-linearities the solution is still dominated by the term, $A\cos\omega t$, while the frequency-amplitude relationship is altered by a slight increase in the coefficient of the term $\beta A^2$. The discussion leading to Eq. (8) indicates that the upper limit of this coefficient is one.

If only the first two terms of Eq. (18) are retained, and if the maximum value of $\theta$ is designated with a subscript $m$, then Eq. (19) can be written

$$\omega_2^2 = \omega_0^2 + 0.71\beta\theta_{2m}^2 - F/\theta_{2m}. \quad (20)$$

Under similar circumstances Eq. (13) can be written

$$\omega_1^2 = \omega_0^2 + 0.69\beta\theta_{1m}^2 - F/\theta_{1m}. \quad (21)$$

By comparing Eqs. (20) and (21) we see that as the iteration process is repeated the coefficient of the term $\beta\theta_m^2$ does increase slightly. The applied physicist is therefore faced with the task of repeating the process until this coefficient is determined accurately enough to give satisfactory results for the particular non-linearity involved. It seemed to this author much more appealing to determine the coefficient of $\beta A^2$ from experimental results.

The results of many experiments indicates that for systems having non-linearities between 0.1 and $\infty$, the value of 0.9 is a good approximation for the coefficient of the term $\beta\theta_m^2$. For specific results see Table I. In Fig. 2 we see the actual experimental values of amplitude and frequency obtained for three typical non-linear systems with different amounts of non-linearity. In these three examples the term $F/\theta_m$ was negligible. In Table II we see the

TABLE III. Typical experimental results showing that the relationship between the amplitude $\theta_m$, and the frequency $\omega$, is essentially independent of the small damping torque $c\dot{\theta}$.*

| $\theta_m$ in millirad. Constants | → $c\times 10^4$ ↓ | 6.4 | 12.9 | 19.3 | 25.8 | 32.1 | 38.6 | 45.1 |
|---|---|---|---|---|---|---|---|---|
| | | | | $\omega$ in radians per second | | | | |
| | 1.05 | 10.2 | 12.9 | 16.4 | 20.9 | 25.3 | 29.8 | 33.8 |
| | 2.39 | 10.2 | 13.2 | 16.9 | 21.2 | 25.8 | 30.2 | 34.2 |
| $k_1 = 0.15 \times 10^7$ | 2.64 | 10.5 | 13.3 | 16.9 | 21.3 | 25.8 | 29.8 | 33.9 |
| $k_3 = 9.7 \times 10^9$ | 3.04 | 10.7 | 13.3 | 17.3 | 21.3 | 26.8 | 30.2 | 34.2 |
| | 3.46 | 10.7 | 13.8 | 17.3 | 21.3 | 26.0 | 30.2 | 34.2 |
| $k = 32.6$ | | | | | | | | |
| $k_1 = 0.10 \times 10^7$ | 1.05 | | 12.4 | 16.9 | 21.8 | 26.3 | 31.0 | 35.3 |
| $k_3 = 10.2 \times 10^9$ | 1.95 | 9.8 | 12.9 | 16.9 | 21.8 | 26.5 | 31.1 | 35.4 |
| $k = 32.6$ | 2.39 | 9.8 | 12.9 | 16.9 | 21.8 | 26.7 | 31.1 | 35.6 |
| | 2.64 | 9.8 | 12.9 | 17.6 | 21.8 | 26.4 | 31.1 | 35.6 |
| | 3.04 | 9.8 | 13.2 | 17.4 | 22.2 | 26.7 | 31.1 | 35.6 |
| | 3.46 | 10.2 | 13.3 | 17.8 | 22.2 | 26.9 | 31.6 | 35.7 |

* The magnitude of the applied torque is $k\omega^2$.

TABLE IV. Typical phase angle data for non-linear systems.*

| Spring constants | Amplitude millirad. | Frequency rad./sec. | Phase angle degrees | Non-linearity |
|---|---|---|---|---|
| | 3.2 | 24.9 | 17 | 0.01 |
| $k_1=1.48\times10^7$ | 6.4 | 27.9 | 20 | 0.03 |
| $k_3=11.08\times10^9$ | 14.2 | 29.2 | 27 | 0.15 |
| | 19.3 | 33.1 | 34 | 0.28 |
| | 25.8 | 36.3 | 35 | 0.50 |
| | 32.8 | 39.2 | 38 | 0.80 |
| | 45.0 | 44.8 | 52 | 1.51 |
| | ★ *R* | ★ *R* | 90 | |
| | 3.9 | 15.1 | 35 | 0.03 |
| $k_1=0.60\times10^7$ | 7.1 | 17.8 | 50 | 0.09 |
| $k_3=11.08\times10^9$ | 12.9 | 20.1 | 49 | 0.30 |
| | 19.7 | 23.5 | 40 | 0.71 |
| | 26.7 | 27.6 | 46 | 1.38 |
| | 32.2 | 30.8 | 50 | 1.92 |
| | 35.4 | 32.9 | 53 | 2.31 |
| | ★ *R* | ★ *R* | | |
| | 20.3 | 17.8 | 78 | 3.51 |
| $k_1=0.13\times10^7$ | 26.1 | 21.5 | 70 | 5.80 |
| $k_3=11.08\times10^9$ | 30.0 | 25.1 | 72 | 7.67 |
| | 33.2 | 27.0 | 69 | 9.40 |
| | 35.4 | 29.0 | 69 | 10.70 |
| | 38.0 | 31.1 | 72 | 12.31 |
| | 41.8 | 33.5 | 82 | 14.88 |
| | 43.8 | 34.5 | 82 | 16.35 |
| | 45.1*R* | 35.2*R* | 90 | 17.30 |

* Note: (1) *R* indicates resonance. (2) ★ indicates off scale.

results of the harmonic analysis of the wave forms of several typical solutions. We note that even for large non-linearity, the fundamental is the dominating term in the solution.

Consider next the effect of introducing a small amount of viscous damping into the system under consideration. An examination of reference 3 shows that the methods therein considered are suitable for systems with small non-linearities. For small damping the methods yield a frequency-amplitude relationship essentially the same as Eq. (9) except at resonance. At resonance the methods tell us that the phase angle must be approximately 90° and the frequency and amplitude must, in addition to Eq. (9), satisfy the relationship

$$A_r=F/c\omega, \tag{22}$$

in which $A_r$ is the resonance amplitude, $F$ is the magnitude of the applied force, $c$ is the coefficient (small) of viscous damping, and $\omega$ is the frequency. We note that Eq. (22) is the usual approximation for the linear case when the damping is small; however, $\omega$ is now a function of $A$. See Eq. (100.30), reference 3.

Turning now to reference 4, p. 42, we find that although the procedure is different the first approximation yields the same result. For example, consider a system represented by the differential equation,

$$\ddot{\theta}+c\dot{\theta}+\omega_0^2\theta+\beta\theta^3=F\cos\omega t-G\sin\omega t. \tag{23}$$

Upon assuming an initial solution, $\theta_0=A\cos\omega t$, and imposing the conditions of periodicity, we find

$$cA\omega=G, \tag{24}$$

and

$$(\omega_0^2-\omega^2)A+3\beta A^3/4=F. \tag{25}$$

Equations (24) and (25) can be combined in the form

$$S_1^2(\omega, A)+c^2A^2\omega^2=H^2, \tag{26}$$

TABLE V. A comparison of the experimental resonance amplitude $\theta_r$, and the theoretical amplitude $F/c\omega$, for various non-linearities.

| Spring identification | $F/c\omega$ millirad. | Experimental $\theta_r$ (millirad.) | $L=c\omega\theta_r/F$ | Non-linearity |
|---|---|---|---|---|
| 13;0–0 | 20.0 | 28.2 | 1.41 | 0.95 |
| | 15.8 | 20.7 | 1.31 | 0.51 |
| | 12.5 | 14.2 | 1.14 | 0.24 |
| | 10.6 | 11.5 | 1.09 | 0.16 |
| 13;0–1 | 18.1 | 24.8 | 1.37 | 0.74 |
| | 14.6 | 18.3 | 1.25 | 0.40 |
| | 11.8 | 13.3 | 1.13 | 0.21 |
| | 10.1 | 10.7 | 1.06 | 0.14 |
| 13;0–2 | 17.1 | 23.1 | 1.35 | 0.66 |
| | 14.0 | 17.0 | 1.21 | 0.36 |
| | 11.5 | 12.7 | 1.10 | 0.20 |
| | 9.9 | 10.4 | 1.05 | 0.13 |
| 13;0–3 | 16.1 | 21.7 | 1.35 | 0.51 |
| | 13.5 | 16.0 | 1.19 | 0.28 |
| | 11.2 | 12.4 | 1.11 | 0.17 |
| | 9.7 | 10.0 | 1.03 | 0.11 |
| 13;0–4 | 15.5 | 20.6 | 1.33 | 0.48 |
| | 13.1 | 15.3 | 1.17 | 0.26 |
| | 11.0 | 12.0 | 1.09 | 0.16 |
| | 9.5 | 9.8 | 1.03 | 0.11 |
| 13;0–5 | 14.6 | 19.3 | 1.32 | 0.40 |
| | 12.7 | 14.6 | 1.15 | 0.23 |
| | 10.7 | 11.7 | 1.09 | 0.15 |
| | 9.3 | 9.6 | 1.03 | 0.10 |
| 13;0–6 | 14.4 | 17.7 | 1.23 | 0.31 |
| | 12.3 | 13.8 | 1.12 | 0.19 |
| | 10.4 | 11.3 | 1.09 | 0.13 |
| | 9.0 | 9.3 | 1.03 | 0.09 |
| 13;0–7 | 13.9 | 16.9 | 1.22 | 0.31 |
| | 12.1 | 13.4 | 1.11 | 0.19 |
| | 10.2 | 11.0 | 1.08 | 0.13 |
| | 8.8 | 9.1 | 1.03 | 0.09 |
| 13;0–8 | 13.6 | 16.1 | 1.18 | 0.27 |
| | 11.8 | 12.9 | 1.09 | 0.17 |
| | 10.1 | 10.8 | 1.07 | 0.12 |
| | 8.8 | 9.0 | 1.02 | 0.08 |
| 20;0–2*B*94 | 23.7 | 45.1 | 1.90 | 13.4 |
| | 21.6 | 38.5 | 1.78 | 9.8 |
| | 19.0 | 32.2 | 1.70 | 6.8 |
| | 16.4 | 25.8 | 1.57 | 4.4 |
| | 15.1 | 19.3 | 1.28 | 2.5 |
| 20;0–2*B*93 | 23.0 | 36.4 | 1.58 | 2.1 |
| | 19.7 | 30.3 | 1.54 | 1.5 |
| | 16.3 | 24.2 | 1.48 | 0.9 |
| | 13.6 | 17.7 | 1.30 | 0.5 |
| | 9.3 | 10.3 | 1.11 | 0.2 |

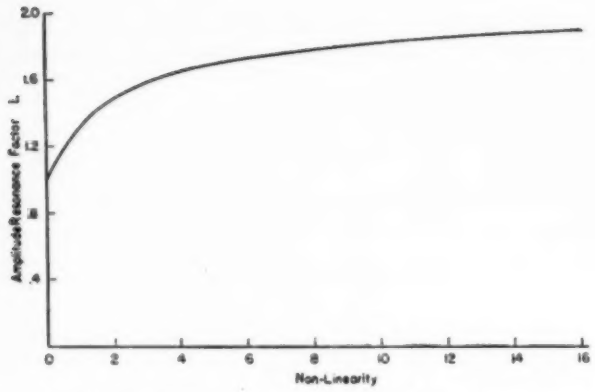


FIG. 3. The factor $L$ as a function of the non-linearity.

in which

$$S_1(\omega, A) = (\omega_0^2 - \omega^2)A + 3/4\beta A^3, \quad (27)$$

and

$$H^2 = F^2 + G^2. \quad (28)$$

At resonance $S_1(\omega, A) = c^2A/2$, $H = F$, and for small values of $c$, Eq. (26) becomes

$$A_r = F/c\omega, \quad (29)$$

which is the same as Eq. (22). It is perhaps well to mention at this time, that in non-linear systems the resonance curve is discontinuous and exhibits the so-called "jump" phenomena. These discontinuities or jumps occur at different frequencies depending upon whether the frequency is being increased or decreased. In the system described by Eq. (23) with $\beta$-positive, the larger resonance amplitude occurs when the frequency is being increased. Throughout this paper this is the resonance amplitude referred to.

If we repeat the process assuming a solution

$$\theta_1 = A\cos\omega t + (\beta A^3/32\omega_1^2)\cos\omega t, \quad (30)$$

Eq. (24) remains unchanged, and Eq. (25) becomes

$$(\omega_0^2 - \omega^2)A + (3\beta A^3/4) \times [1 + \beta A^2/32\omega_1^2 + \beta^2 A^4/512\omega_1^4] = F \quad (31)$$

in which $\omega_1^2$ is given by Eq. (13). If we consider now systems with non-linearity so large that Eq. (16) can replace Eq. (13) then Eq. (31) becomes

$$(\omega_0^2 - \omega^2)A + (3\beta/4)(397/384)A^3 = F. \quad (32)$$

Upon squaring and adding Eqs. (24) and (32) we have

$$S_2^2(\omega, A) + c^2A^2\omega^2 = H^2, \quad (33)$$

in which

$$S_2(\omega, A) = (\omega_0^2 - \omega^2)A + (3\beta/4)(397/384)A^3. \quad (34)$$

Proceeding as before, $S_2(\omega, A) = c^2A/2$ at resonance, and for small values of $c$, Eq. (33) becomes,

$$A_r = F/c\omega. \quad (35)$$

A comparison of Eqs. (29) and (35) shows that no matter how many times we repeat this process, the amplitude of the fundamental at resonance will always be given by the same equation utilizing however higher and higher approximations for the frequency. In terms of resonance amplitude of the solution rather than of the fundamental, Eqs. (29) and (35) become

$$\theta_{1r} = F/c\omega, \quad (36)$$

TABLE VI. A comparison of the ratio $\omega/\omega_0$ with $L$, the corrective factor in the equation $\theta_r = LF/c\omega$.

| Spring identification | $\omega/\omega_0$ | $L$ | Non-linearity |
|---|---|---|---|
| 13;0–0 | 1.40 | 1.41 | 0.95 |
| | 1.22 | 1.31 | 0.51 |
| | 1.11 | 1.14 | 0.24 |
| | 1.07 | | |
| 13;0–1 | 1.32 | 1.37 | 0.74 |
| | 1.18 | 1.25 | 0.40 |
| | 1.10 | 1.13 | 0.21 |
| | 1.07 | 1.06 | 0.14 |
| 13;0–2 | 1.27 | 1.35 | 0.66 |
| | 1.15 | 1.21 | 0.36 |
| | 1.08 | 1.10 | 0.20 |
| | 1.06 | 1.05 | 0.13 |
| 13;0–3 | 1.22 | 1.35 | 0.51 |
| | 1.13 | 1.19 | 0.28 |
| | 1.08 | 1.11 | 0.17 |
| | 1.06 | 1.03 | 0.11 |
| 13;0–4 | 1.20 | 1.33 | 0.48 |
| | 1.11 | 1.17 | 0.26 |
| | 1.07 | 1.09 | 0.16 |
| | 1.05 | 1.03 | 0.11 |
| 13;0–5 | 1.14 | 1.32 | 0.40 |
| | 1.09 | 1.15 | 0.23 |
| | 1.06 | 1.09 | 0.15 |
| | 1.04 | 1.03 | 0.10 |
| 13;0–6 | 1.15 | 1.23 | 0.31 |
| | 1.09 | 1.12 | 0.19 |
| | 1.06 | 1.09 | 0.13 |
| | 1.04 | 1.03 | 0.09 |
| 13;0–7 | 1.14 | 1.22 | 0.31 |
| | 1.09 | 1.11 | 0.19 |
| | 1.06 | 1.08 | 0.13 |
| | 1.04 | 1.03 | 0.09 |
| 13;0–8 | 1.12 | 1.18 | 0.27 |
| | 1.08 | 1.09 | 0.17 |
| | 1.06 | 1.07 | 0.12 |
| | 1.04 | 1.02 | 0.08 |
| 20;0–2*B*94 | 3.47 | 1.90 | 13.4 |
| | 3.07 | 1.78 | 9.8 |
| | 2.66 | 1.70 | 6.8 |
| | 2.20 | 1.57 | 4.4 |
| | 1.76 | 1.28 | 2.5 |
| | 1.37 | 1.09 | 1.2 |
| 20;0–2*B*93 | 1.86 | 1.58 | 2.1 |
| | 1.67 | 1.54 | 1.5 |
| | 1.48 | 1.48 | 0.9 |
| | 1.32 | 1.30 | 0.5 |
| | 1.14 | 1.11 | 0.2 |

and

$$\theta_{2r} = (25/24)F/c\omega. \tag{37}$$

A comparison of Eqs. (36) and (37) reveals that for systems with large non-linearities the amplitude at resonance will be larger than that predicted by the theory of the first approximation.

Thus for large non-linearity we have the following conclusions concerning periodic solutions of Eq. (23) which have a frequency the same as the frequency of the forcing function:

(1) For most practical purposes, the solution can be considered sinusoidal in form.

(2) The frequency-amplitude relationship for all large non-linearities is approximately

$$\omega^2 = \omega_0^2 + 0.9\beta\theta_m^2 - F/\theta_m, \tag{38}$$

and is essentially independent of damping except at resonance.

(3) At resonance not only must Eq. (38) be satisfied but also the relationship

$$\theta_r = LH/c\omega, \quad L \geqslant 1 \tag{39}$$

must hold true. In Eq. (39) the value of $L$ depends upon the value of the non-linearity; $L$ approaches 1 as the non-linearity approaches zero.

As with the coefficient of the term $\beta\theta_m^2$, this author decided to determine the relationship between $L$ and the non-linearity from experimental results.

## EXPERIMENTAL RESULTS

Experimental results have already been used to verify our first two conclusions; and some additional data is displayed in Table III to show that the frequency-amplitude relationship is essentially independent of small amounts of viscous damping except at resonance.

For the case of resonance we note from Table IV that, as expected, the phase angle between the applied force and the solution is approximately 90° at resonance if the damping is small. The dependence of the factor $L$ on the non-linearity $k_3\theta_r^2/k_1$ is shown in Table V. We note that $L>1$, $L\to 1$ as the non-linearity approaches zero, and $L$ increases as the non-linearity increases. These are just the results which can be inferred from Eqs. (36) and (37). A plot of $L$ *versus* non-linearity is shown in Fig. 3, and supplies the information needed to make Eq. (39) valuable in the actual solution of problems.

It is interesting to note at this point that Eq. (100.23) p. 316 of reference 3 indicates that for not too large non-linearity the value of $L$ should be $\omega/\omega_0$ if the forcing function is supplied by a rotating unbalance. A comparison of $L$ and $\omega/\omega_0$ is made in Table VI and we note that they are approximately equal provided the non-linearity remains less than one.

## EQUIVALENT LINEARIZATION

Returning to Fig. 3, we note that the results of this figure together with Eqs. (38) and (39) are sufficient to determine $\theta_r$. The procedure is simply to choose a value of $\theta_r$, compute the non-linearity $k_3\theta_r^2/k_1$, and with this value of non-linearity find the value of $L$ from Fig. 3. Using this value of $L$ and the chosen value of $\theta_r$, solve Eq. (39) for $\omega$. This value of $\omega$ and the original choice of $\theta_r$ must then satisfy Eq. (38). When this is true, the choice of $\theta_r$ is correct.

It is obvious that this procedure, which resulted from a study of Eq. (23) can be generalized to include systems with small viscous damping and any kind of a non-linear restoring force. It should be noted, however, that the periodic solutions are restricted to have the same frequency as the forcing function. These are not the only periodic solutions to equations of this type and the author will report on his investigations on subharmonic solutions in a subsequent paper.

Since for most practical purposes, the solutions examined in this paper can be considered sinusoidal, and since Eq. (39) differs from the resonance amplitude equation for a linear system only by the factor $L$, the procedure outlined above is essentially a method of equivalent linearization for systems with large non-linearity; the equivalent linearization of $\omega$ being defined by Eq. (38).

## ACKNOWLEDGMENTS

This research was made possible by a contract with the David Taylor Model Basin, U. S. Navy Department; and a Frederick Gardner Cottrell grant-in-aid from the Research Corporation.

# The Illuminating System of the Electron Microscope

JAMES HILLIER AND S. G. ELLIS
*RCA Laboratories, Princeton, New Jersey*
(Received December 2, 1948)

Using a self-biased electron gun the current arriving at the specimen, the current density at the specimen, and the area of illumination of the specimen have been determined for different values of the filament height. The variation of current density at the specimen with angular aperture of illumination has been determined and the results compared with those for the zero-biased gun. The performance of the self-biased gun is assessed and some of the factors involved in its operation and maintenance are discussed.

## INTRODUCTION

THE purpose of this paper is to provide data on the illuminating system of an electron microscope when it is operated with an electron gun biased with a cathode resistor and operated at saturated beam current. For simplicity in terminology such an electron gun operated at saturated beam current will be called a self-biased gun. The paper also contains a discussion of some of the factors to be taken into consideration in the operation of such an illuminating system.

Hillier and Baker[1] have described the adjustment and performance of an illuminating system employing a fixed-bias gun with zero bias. Ellis[2] has described the optics of three-electrode electron guns and has also described qualitatively the behavior of the self-biased gun. References to the earlier literature will be found in the above papers.

For efficient operation the illuminating system should satisfy the following requirements. First, it should illuminate only that part of the specimen which is under examination, for illumination which reaches other parts of the specimen serves only to heat the specimen. The temperature of the region under examination and therefore the tendency to specimen destruction is increased without any gain of image intensity.[3,4] The general heating also tends to cause thermal drifting of the specimen. The illumination of parts of the specimen that are not under observation has the further undesirable effects that these parts of the specimen are contaminated, and also that scattered electrons from them may enter the objective lens and decrease the contrast in the final image. The electron gun should therefore provide a small and preferably circular source. Secondly, the illuminating system should provide sufficient intensity of illumination at the desired angular aperture that the microscope may be readily focused. Thirdly, the angular aperture of the illumination should, for most purposes, be as small as possible, consistent with the foregoing requirements. Small angular aperture of the illumination is desirable for two reasons: if the objective lens has any asymmetries it is better to use the more paraxial portions of the lens. This minimizes the effects of the asymmetries which become more apparent as the rays traversing the objective do so further from the axis. The other reason for requiring small angular aperture of illumination is that it will tend to sharpen and make more distinct the contour and phase phenomena in the final image and hence facilitate the detection of exact focus. It is now possible and indeed desirable to sacrifice intensity in order to achieve small angular aperture of illumination since the focusing can be performed at lower intensity of illumination if a high magnification viewer is employed either with or without a deflection focusing device of the type described by le Poole[5] and modified by the authors. The total beam current reaching the specimen should be kept as small as possible, consistent with the foregoing requirements, since for all the condenser strengths employed in practice this beam current, apart from the fraction hitting the specimen supporting screen, passes into the objective lens. If an objective diaphragm is employed, that part of the beam current which does not go through the diaphragm deposits contamination upon it. When the beam current is large the rate of contamination is large and moreover the static charge maintained by the beam current on the diaphragm is considerable so that the fraction of the beam current which proceeds through the diaphragm to form the intermediate image suffers irregular deflections as a result of the electrostatic field produced by this charge. Even if no objective diaphragm is used, the rate of contamination of the objective lens will increase as the current reaching the specimen is made larger. The foregoing optical requirements cannot be met independently and are, in fact, to a certain extent contradictory. In practice a compromise is reached by producing a small single intense source in the electron gun and by choosing the best condenser

[1] J. Hillier and R. F. Baker, J. App. Phys. **16**, 469 (1945).
[2] S. G. Ellis, J. App. Phys. **18**, 879 (1947).
[3] M. von Ardenne, Kolloid Zeits. **108**, 195 (1944).
[4] V. K. Zworykin and J. Hillier, J. App. Phys. **14**, 658 (1943).
[5] J. B. le Poole, Philips Tech. Rev. **9**, 33 (1947).

diaphragm radius and condenser lens strength for the immediate purpose in view.

In addition to the optical requirements, certain mechanical requirements have to be met. These are briefly that the illuminating system shall be as simple in construction as possible, that it shall not be too difficult to align it to the objective lens, and that the illuminating system shall be readily demountable for the renewal of the filament and the cleaning of the whole system. Finally, the operator may in special cases, wish to employ unusually high current densities of bombardment of the specimen or other unusual conditions of operation. It is, therefore, convenient if possible to avoid making the design so rigid in concept as to preclude such special departures from normal practice.

The measurements reported below were made on the RCA model EMU electron microscope in use in this Laboratory. The symbols employed in this paper are those previously introduced in references 1 and 2. They are listed below, together with the parameters of the illuminating system of this microscope.

$u_k$ = spacing of filament above lower surface of the gun focusing aperture.
$u_c$ = height of effective source above the condenser lens center.
$v_c$ = distance from condenser lens center to the specimen = 150 mm.
$2r_k''$ = diameter of focusing aperture = 3.75 mm (0.15 in.).
$2r_c''$ = diameter of condenser aperture = 0.5 mm (0.020 in.).
$2r_o'$ = diameter of region in specimen plane illuminated when the intensity is a maximum, (i.e., focused illumination).
$f_c$ = focal length of condenser lens.
$I_c$ = condenser lens current (in test meter scale divisions unless otherwise stated).
$I_B$ = beam current (in beam current meter reading minus leakage current shown when the filament is off).
$i_B$ = current reaching the specimen plane.
$\sigma$ = current density in beam reaching the specimen plane.
$\sigma_0$ = maximum value of $\sigma$.
$\eta' = \sigma_0/I_B$ the illuminating efficiency of the illuminating system.
$\alpha_c$ = aperture of the illumination at the specimen.

The shield to anode potential difference was 50 kilovolts. The filament diameter employed was 4 mils ($=10^{-2}$ cm), the focusing aperture to anode spacing was 15 mm (approx.), and the anode aperture 18 mm. The biasing resistor was 0.5 megohm.

## EXPERIMENTAL PROCEDURE

With the condenser lens strength adjusted to provide maximum illumination, the following quantities were measured: the total current arriving at the specimen, the distribution of current density in the illuminating beam at the specimen, and the size and shape of the area of illumination at the specimen. These measurements were obtained for different values of the filament-aperture spacing using the self-biased gun. For comparison purposes similar measurements were made on the zero bias gun by shorting out the cathode biasing resistor and adjusting the filament temperature to produce the same total beam current leaving the filament.

The apparatus used in making these measurements may be described with the aid of Fig. 1. A collector plate $E$ about 1 square centimeter in extent was mounted on the axis of the microscope, midway between the projector lens and the final viewing screen. It was maintained at a potential of one hundred volts positive with respect to the body of the instrument by means of the battery $B$ and the current reaching it was measured by a sensitive electronic microammeter $M$. Parallel plates $D_1$, $D_2$ were mounted between the objective and condenser lenses so that the illuminating beam could be deflected through small angles.

The total beam current was measured by removing the projector polepieces and focusing the microscope so that the current passing through the specimen plane was completely collected by the plate $E$. In measuring the distribution of current density at the specimen plane the microscope was set at normal focus and operated at a magnification of approximately 6000. The microscope was aligned so that the microammeter gave a maximum reading and then voltages were applied to the plates $D_1$ and $D_2$ so that the collector $E$ received, successively, currents from areas spaced out radially from the center of the illuminated region. The filament was oriented in the electron gun so that the deflection produced by the plates $D_1$, $D_2$ was parallel to the short diameter of the area of illumination. The collector $E$ was conjugate to an area of approximately $3\times3$ microns in the normal specimen plane. The measurements of distribution of current density were, therefore, somewhat crude. To obtain the size and shape of the area of illumination of the specimen, the microscope magnification was changed to

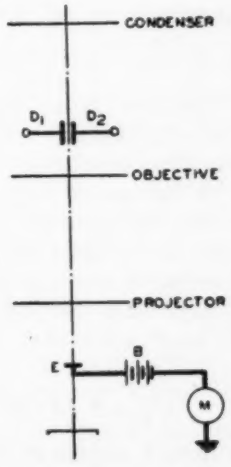


FIG. 1.

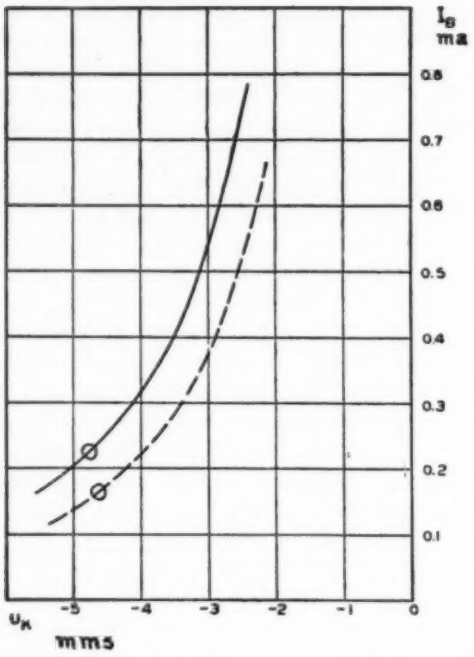


FIG. 2. Variation of total beam current with filament height. Full line experimental self-biased gun used in these experiments ($r_k''=0.150$ inch). Dashed line standard RCA self-biased gun for comparison ($r_k''=0.125$ inch).

approximately 1000 and the spot of illumination photographed with exposures of 1, 2, 4, 8, and 16 seconds. Traces from these plates were obtained on a Leeds and Northrup recording microphotometer and from these traces, it was possible to determine the distribution in intensity and also the half-width of the area of illumination along the long and short diameters. In order to be able to determine the intensities of illumination under the conditions in which photography is normally carried out, the relation between condenser lens current and lens strength was determined in the following way. The radius $r_0'$ of the area of illumination at the specimen was measured for different condenser lens currents $I_c$ and the condenser lens strength $1/f_c$ deduced from the approximate relation

$$1/f_c=1/v_c[(2r_o'/r_c'')+1]-1/u_c,$$

where $u_c$=distance from condenser lens to the source in the electron gun=160 mm with the self-biased gun. The condenser lens current $I_c$ was measured on the panel test meter which has 100 scale divisions corresponding to 25-ma full scale. The relations and graphs given below are in terms of scale divisions in order to simplify their use.

It was found that the condenser lens strength was given very accurately by Busch's formula where

$$1/f_c=2\cdot51\times10^{-5}\cdot I_c^2,$$

and $f_c$ is in mm, $I_c$ in test meter scale divisions.

## EXPERIMENTAL RESULTS

In presenting the experimental results some consideration must be given to the choice of independent variable. The filament to focusing aperture distance $u_k$ is the quantity which is preset and controls the behavior of both the self-biased and the zero bias guns. When operating the self-biased gun the total beam current leaving the gun is the quantity most readily observed in the laboratory. The relation between these quantities is shown in Fig. 2. In operating the zero bias gun the optimum filament height is ascertained by observations of the nature of the illumination at the specimen.[1] With this gun the filament height controls the distance of the effective source from the condenser lens, and consequently affects the condenser current required to focus an image of the source at the specimen, Fig. 3a. It will be seen from Fig. 3b that within the range of filament heights normally employed the self-biased gun produces an effective source at a nearly constant distance from the condenser lens.

In the following sections the filament height will be treated as the independent variable. The variation of the beam current $i_B$ arriving at the specimen with filament height is given in Fig. 4. In Fig. 5 the current density, $\sigma_0$ at the center of the focused illumination is shown as a function of filament height. It may be mentioned in passing, that the focusing is quite critical. Moreover the maximum current density attainable is dependent upon the

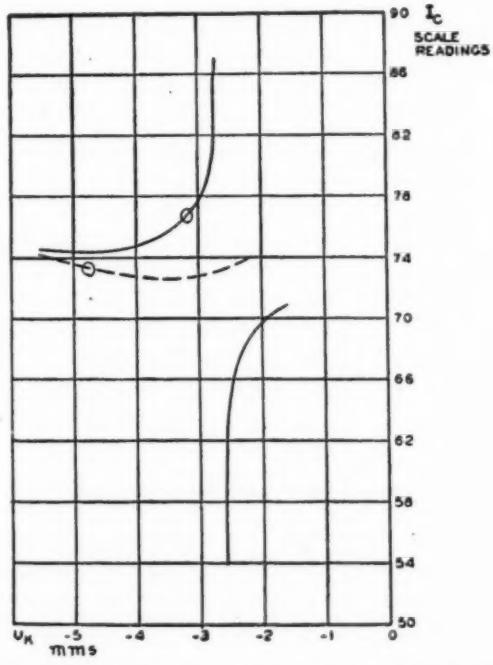


FIG. 3a. Condenser current required to focus illumination for different filament heights. Full line zero bias gun, dashed line self-biased gun.

shape of the filament being larger for filaments which are sharply bent at their apexes.

The illuminating efficiency defined as $\sigma_0/I_B$ (see Fig. 6) is of interest because a high efficiency indicates that good illumination can be obtained with reasonably low values of $I_B$, which in turn implies that a lower filament temperature, giving a longer filament life, can be used than if the efficiency were poor. This argument is not complete. The self-biased gun is operating closer to cut-off so that somewhat higher filament temperatures are required to yield the same total beam current $I_B$. The filament of the self-biased gun is usually run at a higher temperature than that of the zero bias gun, but the difference in temperature and life of the filament is not large enough to be objectionable in view of the higher intensity of illumination obtained. It is here assumed that the filament temperature is not raised above the value required to produce saturation of the beam current.

Since for visual focusing it is desirable to obtain good illumination of that part of the specimen which is being observed, and at the same time have small total illumination of the specimen to minimize contamination the variation of the quantity $\sigma_0/i_B$ with filament height is of interest. This is shown in Fig. 7. It may be concluded that it is desirable to operate the self-biased gun with a filament height of 4.8 mm and choose the condenser diaphragm radius to provide the proper level of maximum illumination for the work in hand. For the standard RCA biased gun the optimum filament height will be slightly less.

The quantity $\sigma_0/i_B$ is obviously connected with

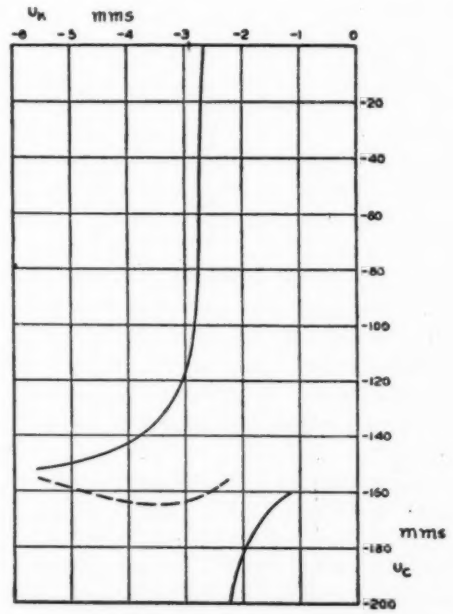


FIG. 3b. Effective source positions relative to condenser lens. Full line zero bias gun, dashed line self-biased gun.

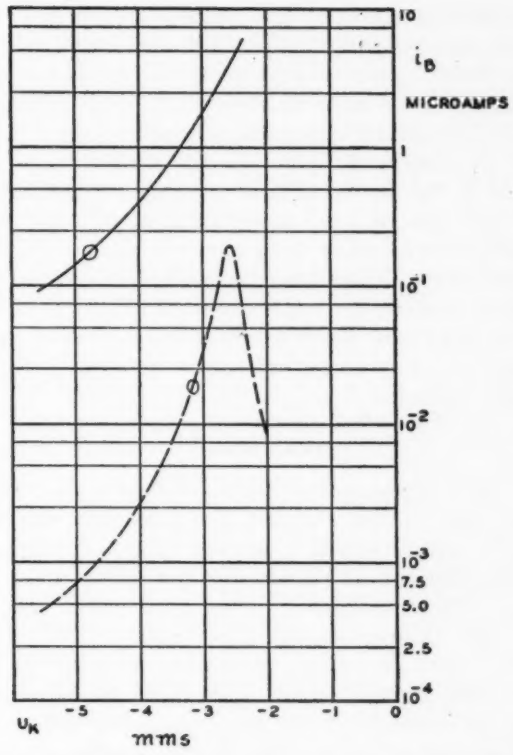


FIG. 4. Current reaching specimen for different filament heights. Full line self-biased gun, dashed line zero bias gun.

the area of the specimen which is illuminated by the *focused* beam. The relation between $\sigma_0/i_B$ and $r_0'$ the radius at which $\sigma=\sigma_0/2$, is also dependent upon the distribution of intensity in the focused illumination. This distribution is critically dependent on the accuracy of focusing of the illumination. Finally the illuminated area is usually elongated with a length approximately 50 percent greater than the breadth in the best cases when the filament tip is sharply pointed. Figure 8 shows the variation of 'spot radius' $r_o'$, with filament height. The radius recorded here is the minimum radius of the "spot." The points at which the intensity in the illumination fell to half the maximum value were found from photo-densitometer traces. The visual observation of the spot at low magnification on the final viewing screen is not a satisfactory method of determining the spot size since the light output of the fluorescent screen is not a linear function of the incident current density but shows saturation at higher current densities. Visual observation therefore leads to an over estimate of the half-width.

The factors discussed above pertain primarily to the electron gun. With the results there obtained it is a straightforward matter to calculate by the formulae of von Borries and Ruska[6] the dependence of the aperture of illumination $\alpha_e$ and the intensity

[6] B. von Borries and E. Ruska, Zeits. F. Tech. Physik **20**, 225 (1939).

of illumination $\sigma$ upon the condenser lens current $I_c$. These are presented graphically in Figs. 9 and 10. In calculating the values of $\alpha_e$ the minimum spot radius $r_o'$ has been used. Again in Fig. 10 the curves are not shown flat topped. Flat topped curves would only be obtained if the source were circular and had uniform current output across its cross section. The maxima of the curves in Fig. 10 have been adjusted to coincide with the previously observed values of $\sigma_0$. These are in both cases somewhat greater than would have been calculated from the formulas of von Borries and Ruska. This is again due to the way the intensity is distributed in the focused illumination.

## DISCUSSION

In the graphs above, the normal operating points have been circled. It will be seen that the self-biased gun produces a source which is approximately one-half the diameter (or one-quarter the area) of that produced by the zero biased gun. At the same time it provides from ten to fifty times the intensity depending on the filament height and the shape of the filament apex. This means that *at the same intensity* of the final image, focusing and photography can be done at an angular aperture of the illumination which is less, by a factor between five and seven, than that required with the zero bias gun. The total bombardment of the specimen is some twenty times greater with the self-biased gun. When the illumination is focused, the small size of the source tends to minimize the temperature rise in the specimen, the higher intensity is accompanied however by a higher rate of contamination. When the angular aperture is reduced, the heating takes place over a larger area, much of it now in the mesh. The same is true of the contamination, the rate of deposition in the field of view is reduced, but the deposition becomes more widespread.

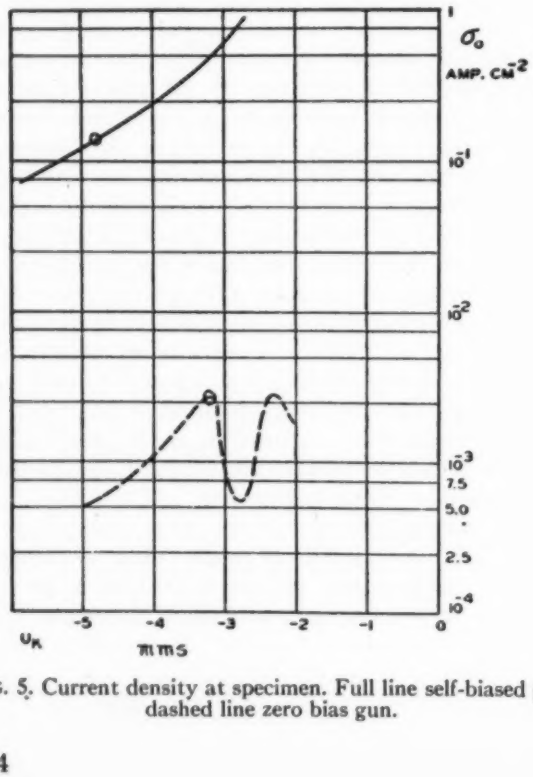


FIG. 5. Current density at specimen. Full line self-biased gun, dashed line zero bias gun.

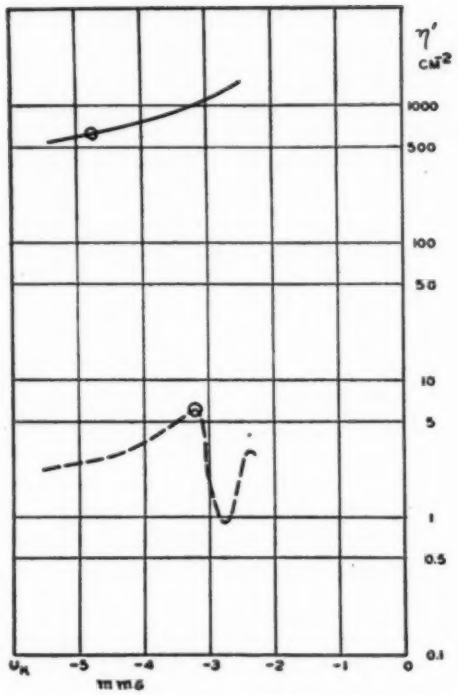


FIG. 6. Illuminating efficiency ($\sigma_0/I_B$). Full line self-biased gun, dashed line zero bias gun.

The above results, together with experience gained in its use, make it possible to assess the general performance of the self-biased gun. Its most notable feature is that it permits focusing and photography at low angular aperture with a consequent sharpening of the contour and phase contrast phenomena. Moreover the high available intensity and inherent crispness of the images facilitate the detection of defects which may occur in the imaging system of the microscope. The next most attractive feature is that the setting of the filament height, filament centering, and gun tilt are not critical. The source is definitely single so that double images are not produced. The single source has great advantages in electron diffraction studies.

For a given filament temperature the optical behavior of the gun will be the same whether the bias is produced by a cathode resistor or a battery. The use of a resistor is obviously simpler, and has the further advantage that saturated beam current operation is attainable. The latter provides a stable electron current through the instrument which results in a stable distribution of charge throughout the imaging system, and a consequent freedom of

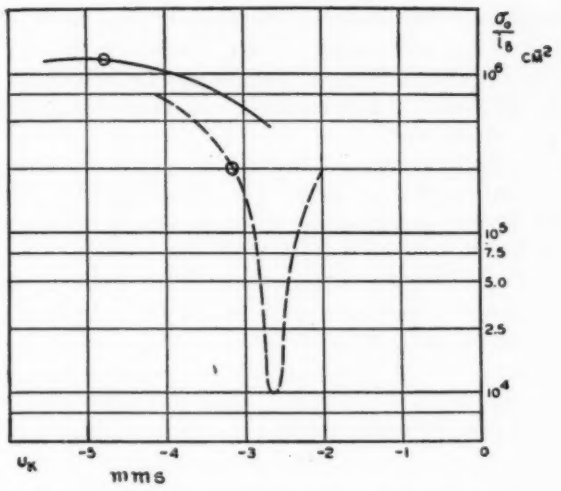


FIG. 7. Full line self-biased gun, dashed line zero bias gun.

the final image from shifts which might otherwise be caused by varying electrostatic deflection of the imaging beam.

The major disadvantages to be cited against the self-biased gun are the heating and contamination caused by the higher current reaching the specimen. The heating effects will be found of negligible importance, however, at the intensities and angular apertures which are proper for focusing. The rate of contamination of the objective lens and its diaphragm can be reduced by using a smaller condenser diaphragm. Finally the higher operating temperature of the filament results in a slightly shorter filament life, the decrease is not great enough to be serious. These disadvantages are far outweighed by the advantages cited above, moreover as indicated their effects can be minimized by the proper handling of the electron microscope.

## THE OPERATION OF A SELF-BIASED GUN

To obtain the best results when using the self-biased gun attention must be given to several factors involved in its operation. The filament tip should be bent as sharply as possible (see Fig. 18 of reference 1). The tip of the filament should be centered with respect to the shield when the filament is installed. Any bending of the filament must be done before it is heated since heating makes the tungsten wire brittle. The behavior of the self-biased gun is not critically dependent on filament centering so that it is sufficient to center the filament by eye.

For stable operation the illuminating system must be clean and must be designed so that insulating surfaces which may charge up cannot deflect the illuminating beam. In the model EMU electron microscope the chief insulating surfaces requiring attention are the fluorescent screen below the anode and the window through which it is viewed. Since in using the self-biased gun this fluorescent screen serves no useful purpose, it is better to clean it down to the metal with steel wool. The wire screen over the viewing window should not be in contact with the window and should be renewed whenever it becomes blackened by contamination. Alternatively the window may be replaced by a metal plate.

With the exception of the condenser polepieces all metal surfaces in the illuminating system should be cleaned with the finest grade steel wool until they shine. This is particularly true of the *inside* of the focusing shield near the aperture, and those other parts of the system such as the condenser aperture insert which are in the direct path of the beam. Strands of steel wool left behind after the cleaning should be carefully removed with a bar magnet. In some places this cleaning will eventually remove the plating. This will not affect the performance of the gun provided that no partially attached pieces of plating are left in the gun. The sharp edges of such plating might lead to field emission during operation.

The insulating surfaces in the electron gun should not be struck by high voltage electrons for in this case they become partially conducting and the gun eventually breaks down. With the self-biased gun it is also desirable to widen the hole through the anode so that electrons do not strike any metallic surface there. If this precaution is not observed a conducting layer may build up on the glass of the electron gun and eventually lead to break down.

The external cleanliness of the electron gun and other surfaces which are at high voltages is im-

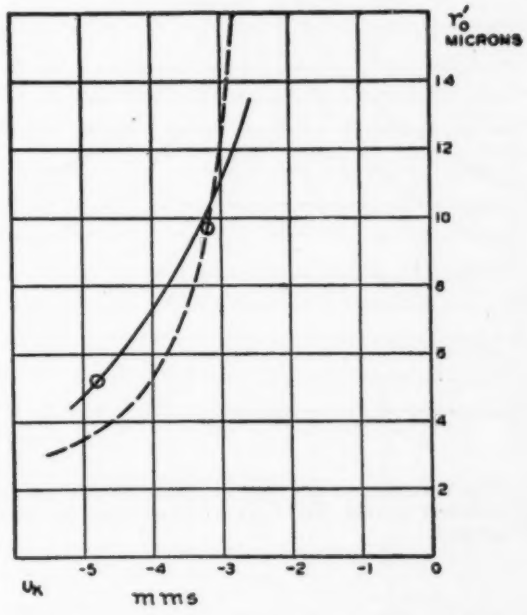


FIG. 8. Spot size variation with filament height. Full line self-biased gun, dashed line zero bias gun.

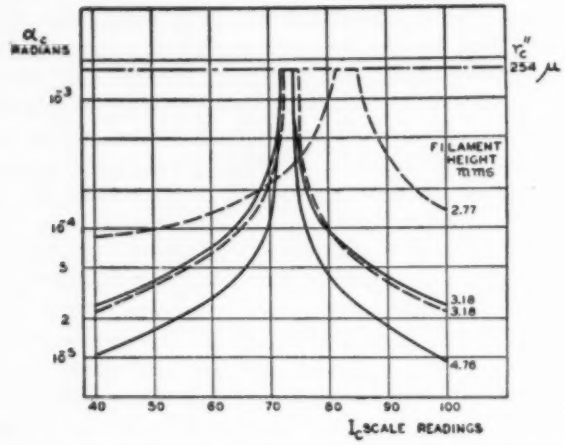


FIG. 9. Variation of angular aperture of illumination with condenser current. Full lines self-biased gun, dashed lines zero bias gun.

portant since dust will cause corona discharges and sometimes spark breakdown. In cleaning the glass of the electron gun care should be taken to avoid damage to the glaze of the external surface. If there is much dust on it, it should be wiped circumferentially rather than longitudinally. Longitudinal scratches in the glass tend to accumulate continuous lines of moisture which serve as a track for breakdown over the surface. Deposits on the inner face of the glass will usually prove intractable if treated with organic solvents. They may be removed by boiling the glass (without the end rings) in a strong caustic solution.

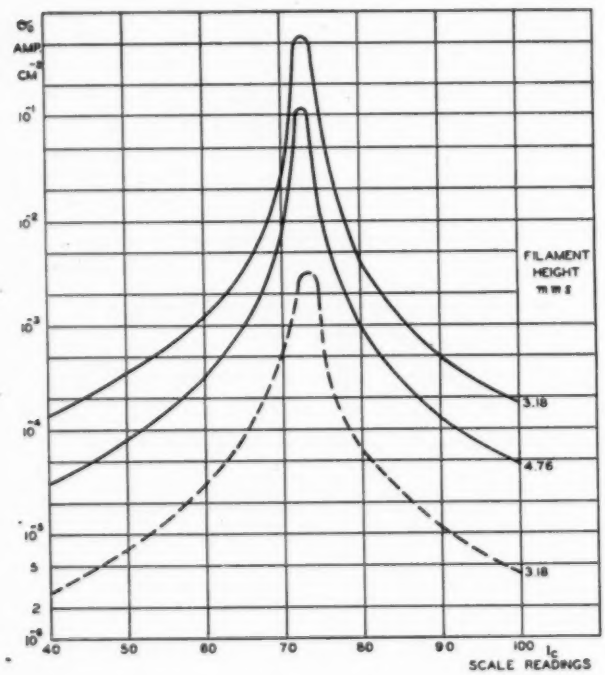


FIG. 10. Variation of intensity of illumination of specimen with condenser current. Full lines self-biased gun, dashed line zero bias gun.

The adjustment of the zero bias electron gun for optimum performance has been described by Hillier and Baker. In the case of the self-biased gun it will be seen from Fig. 5 that the most favorable ratio of intensity of illumination to total current reaching the specimen is obtained with the filament set rather high with respect to the shield. If the filament is set to give a total beam current of 250 microamps at saturation ($-u_k = 4.5$ mm) then the maximum intensity of illumination with a 20-mil diameter aperture in the condenser is of the order 0.1 amp./cm² and the area of illumination approximately 10×16 microns. It is worth repeating that the attainment of high maximum intensity and small source size is dependent on using a filament with a sharp bend at the apex. Residual asymmetries ("wings") on the illumination can usually be removed with the tilt adjustment on the electron gun.

For some samples an intensity of illumination of 0.1 amp./cm² will prove destructive. The operator is not, of course, obliged to peak the illumination. However, if much work is to be done with such samples it is convenient to employ a condenser aperture of smaller diameter so that the maximum illumination will be less intense. When employing a smaller condenser aperture the alignment of the microscope becomes a slower procedure. It is therefore a simplification if good alignment has been previously achieved using a larger aperture and if the condenser lens is not moved when the smaller aperture is inserted. In all cases the condenser aperture should be accurately coaxial in the condenser lens.

If a very large condenser aperture is used, or if the condenser aperture assembly is removed and if the microscope is then focused at low magnification, a relatively large fraction of the total beam current will reach the final viewing screen. Under these conditions the output of x-rays from the final viewing screen may rise above the generally accepted tolerance level. For this reason it is strongly advised that the use of condenser apertures greater than the standard aperture supplied with the instrument be avoided. If the nature of the work involves the use of a larger aperture, the x-ray level should be measured before proceeding with the experiments.

# Point-to-Plane Corona Onsets*

WILLIAM N. ENGLISH** AND LEONARD B. LOEB
*Department of Physics, University of California, Berkeley, California*
(Received December 20, 1948)

The effect of point material and point radius of curvature on positive and negative intermittent corona onset potentials has been studied with a point-to-plane gap in air at atmospheric pressure. The negative Trichel pulse onset strangely is independent of point material but does depend on point history and radius. This surprising result is shown to come from the circumstance that the Trichel pulse onset depends on current densities needed to condition the point surface and yield a higher value of the second Townsend coefficient to give the increased currents. Trichel pulse onset thus does not mark the onset of a self-sustaining discharge. The self-sustaining discharge initiates at lower potentials and leads to currents of a low order until the cathode spot cleans up.

IN a survey article "Recent developments in analysis of the mechanisms of positive and negative coronas in air,"[1] Loeb*** discusses the rather surprising equality of positive and negative corona onset potentials in air, as contrasted with a much lower negative onset in pure $H_2$ and $N_2$. He concludes that this equality must be due to an increase in the work function of the point, corresponding to a decrease in "gamma," the secondary electron emission coefficient, which happens to raise the negative onset to just about the same value as the positive. With this dependence on the "gamma" of the point, one would at first expect that materials of different work function would show different negative corona onset potentials. Indeed, experiments with a water drop point[2] had revealed that here was a substance with such a low "gamma" that apparently negative corona could not take place at all! The purpose of the present work was to look for any solid substances which might exhibit such an effect, and this involved also an investigation into other factors affecting the onsets. A discussion of the nature of the onsets, and of the types of pulses observed, will be found in a previous paper.[3]

## APPARATUS AND TECHNIQUE

The arrangement of the point-to-plane gap is shown in Fig. 1.

The intermittent corona onset potential, $Vg$, was measured by means of a potentiometer calibrated against a standard cell. For positive point, $Vg^+$ was the lowest potential at which burst pulses could be seen on the Dumont 241 oscilloscope, with full gain (100×). For negative point, $Vg^-$ was the lowest potential at which Trichel pulses could be obtained. The observed onset potential for both polarities is considerably affected by the amount of "triggering" ionization which is introduced into the gap in order to start the electron avalanches. A "natural" onset will be defined as that obtained with the ionization from cosmic rays and the usual slightly radioactive surroundings.

Except where otherwise indicated, the gap geometry was as recommended by Loeb,[4] with a ratio of point diameter to gap length of 1/80, and plane radius greater than gap length. The work was done in room air.

## EXPERIMENTAL RESULTS

### Triggering Ionization

The dependence of positive intermittent onset potential on the amount of triggering ionization is marked, since nearly all the negative ions produced in the gap drift into the high field region, and the electrons formed undergo multiplication by collision with gas molecules. Thus, a radioactive source must be used with caution. The practice followed has been to place a weak $\beta$-$\gamma$-ray source (50 grams of thorianite ore) at such a distance as to give a 3- to 5-fold increase in the frequency of burst pulses at onset. This greatly facilitates rapid and accurate determination of $Vg^+$ without lowering it appreciably. Much more triggering than this causes the burst pulses to merge into one another on the oscilloscope screen, and produces a lowering of the onset through cumulative avalanches.

With negative point, unfortunately, the "natural" onset cannot be determined very accurately with

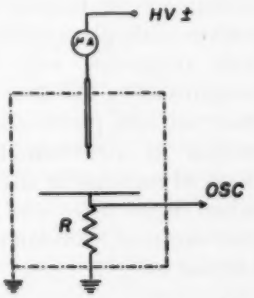


FIG. 1. Diagram of the corona gap, showing the point of insertion of the oscilloscope leads.

* This work was done under contract with the ONR.
** Now at the National Research Council, Chalk River, Canada.
[1] L. B. Loeb, J. App. Phys. **19**, 882 (1948).
*** In a recent article, J. D. Craggs and J. M. Meek, Proc. Phys. Soc. London **61**, 327 (1948).
[2] W. N. English, Phys. Rev. **74**, 179 (1948).
[3] W. N. English, Phys. Rev. **74**, 170 (1948).
[4] J. M. Parker and L. B. Loeb Phys. Rev. **74**, 1557A (1948).

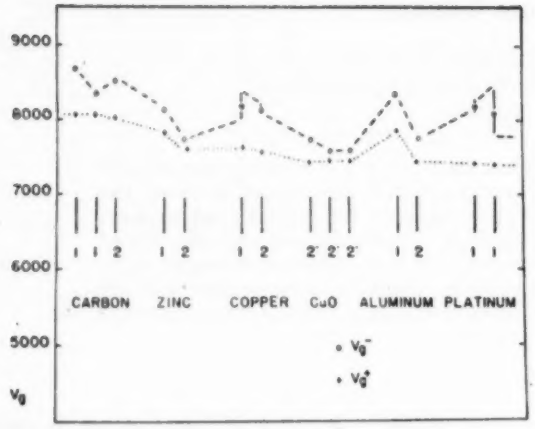


FIG. 2. Characteristic results from the measurements of negative and positive corona onset thresholds, using points made of the materials indicated. The numbers and pointer lines represent different numbered metal points of the series used in the measurement.

the oscilloscope, owing to the long and randomly varying time between successive Trichel pulses. This arises from statistical fluctuations in the rather complex ionizing sequence, from the low and perhaps variable secondary electron emission coefficient of the point, and from the inherently minute sensitive volume for multiplication of electrons, since here the negative carriers move *out* of the high field region, as well as from the limited sweep time of the oscilloscope.

The same weak $\beta$-$\gamma$-ray source mentioned earlier produced an increase in the number of Trichel pulses at onset (and a lowering of the onset) for large negative points (diam. 2 mm), although the effect was of course much less than for the corresponding positive point. It was somewhat surprising to find an appreciable effect also with even the smallest point (diam. 0.03 mm). This was probably a "counting" action due to positive ions from electron tracks. Although it is doubtful if there is much danger of falsifying the negative onset through a cumulative ionization effect, only the minimum amount of triggering necessary to give a reasonably sharp and reproducible onset was used for onset measurements.

An attempt was made to trigger the negative corona by illuminating brass, platinum, and zinc points with ultraviolet light from an Hg arc. Reliable triggering was not obtained owing to the susceptibility of the surfaces to change in work function and photo-emission connected with illumination in air. Sometimes illumination caused a marked increase in the number of pulses at onset; at other times there was little effect, or even a noticeable *decrease*. Similar results have been reported by Loeb *et al.*[5]

[5] Loeb, Kip, Hudson, and Bennett, Phys. Rev. **60**, 714 (1941).

## Point Material

Kip and Bennett[5] found no measurable difference in negative onset for points of platinum, copper, aluminum, iron, and various alloys. In view of the experimental uncertainties in the determination of negative onset, and of the all-important role played by the secondary electron emission coefficient of the point, it was essential to confirm this result, using the positive onset, which is independent of point material, as a check. Accordingly $Vg^+$ and $Vg^-$ were determined (for "natural" ionization), using 1-mm diameter points of carbon, zinc, copper, aluminum, and platinum.

Points were ground hemispherical and carefully polished, but no special effort was made to get identical radii of curvature since the difference between $Vg^+$ and $Vg^-$ gives the information needed. For each determination of the quantity $Vg^- - Vg^+$, four measurements of onset $(+-+-)$ were taken. Some of the results are shown in Fig. 2.

As found in all previous experiments, positive onset was reproducible and apparently independent of point history or material.† The increase in potential, from burst pulse onset $Vg^+$ to the first appearance of streamers, was essentially constant and averaged 270 volts.

Negative onset *as determined* was higher than the

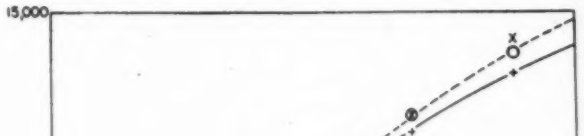


FIG. 3. The onset of positive burst pulse, $Vg^+$, and of negative Trichel pulse corona, $Vg^-$, with and without triggering radioactive sources as a function of point diameter in mm. The decrease in the negative potential relative to the positive onset potential with decreasing radius is clearly seen.

† For example, an onset determination made with the platinum point 9 months earlier checked exactly, although the point had been used considerably in the meantime.

positive by from 125 to 1000 volts.‡ It was notably dependent on point "condition" varying from one measurement to the next, and even during a determination. This was especially true for Pt. Tests showed that the effect was *not* due to changes in the gas filling of the chamber. Taking into consideration the uncertainty in the negative onset, for a given point, there was no significant difference in behavior attributable to point material.

The following side experiments were done:

(1) a A copper point was oxidized black in a bunsen flame.
b Point was etched with dilute nitric acid.
c Point was oxidized again in bunsen flame.

There was no change in positive onset after a, b, or c, and no significant difference in the negative onset, except that the etched point gave more pulses and a cleaner onset.

(2) a A platinum point was bombarded with positive hydrogen ions at 1.5 cm Hg pressure of hydrogen, current 10–20 $\mu$a for 65 minutes.
b $H_2$ was pumped out, air admitted, and negative and positive onsets immediately determined. There was no significant difference from the values previously obtained.

### Point Size

Some early measurements with points of FeS, CuO, and As had given $Vg^-$ considerably below $Vg^+$. These were necessarily rough points because of the nature of the material. The results with the 1-mm points at once suggested that this lowering of $Vg^-$ was due to point shape. Onsets were then determined for a fine steel needle point, a fine hemispherical platinum point (diam. 0.03 mm), and one of the 1-mm carbon points with the tip broken off to give sharp edges. For such fine points the negative onset was *below* the positive onset by several hundred to a thousand volts!

The next step was to study the variation of onset potential with point radius of curvature. A series of 12 polished hemispherically tipped platinum points (diam. 0.03 to 2.4 mm) was made and measured under 100 to 4400 power magnification. The ratio of point diameter to gap length was kept constant at 1/80, as previously indicated. At least three measurements were made for each point, $Vg^-$, $Vg^+$, $Vg^-$, negative onset being determined both for "natural" ionization, and with the radioactive source at the maximum distance which gave a marked increase in pulses near onset. Results are shown in Fig. 3. The consistent character of the data indicates a real and smooth lowering of the Trichel pulse onset relative to the positive onset as point radius decreases.

‡ Similar results have been observed with concentric cylinders by C. G. Miller (see reference 8) in air down to 90 mm pressure.

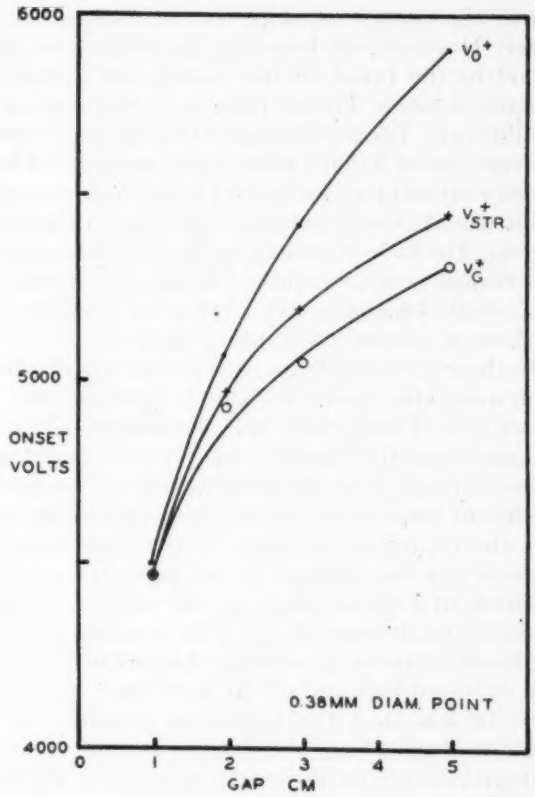


FIG. 4. The various thresholds for positive point corona as a function of the gap length. $Vg^+$ is the beginning of intermittent burst pulse onset and the beginning of the positive point Geiger regime. $V_{str.}^+$ represents the appearance of what are called pre-onset streamers. $V_0^+$ represents the disappearance of streamers and the onset of the continuous burst pulse type corona, marking the end of the Geiger counter regime for the point. It is noted that the counting plateau, which is presumably somewhat shorter than the vertical distance between $Vg^+$ and $V_0^+$, increases with gap length, and that the various thresholds are more clearly delineated the longer the gap for the same diameter point.

### Gap Distance

Although not of fundamental importance, a plot of the three positive onsets *versus* gap length is given in Fig. 4 to show the type of variation. The benefit obtained in using relatively large gaps, such as recommended by Loeb, so that the onsets will be well separated, is evident. Negative onset would be very close to $Vg^+$ for a point of this size, as can be seen from Fig. 3.

## DISCUSSION OF RESULTS

By specifying the level of triggering ionization, in terms of "natural" ionization or otherwise, it is possible to obtain reproducible *positive* onsets from the oscilloscope, although even here the amplification used has some effect on apparent onset. For *negative* onset, a large difference between apparent oscilloscope onset and the "break" in the current-voltage characteristic may be found. Thus, in one case, apparent oscilloscope onset was 900 volts

above the change in slope of the current voltage curve. However, on lowering the voltage to correspond to the break in the curve, and *waiting ten minutes*, a single Trichel pulse was obtained on the oscilloscope. The significance of the abrupt current increase and of Trichel pulse onset is discussed later. A more satisfactory method of triggering a negative point would greatly facilitate the study of negative corona. The lack of success with ultraviolet light in this regard is to be expected from gas-covered and oxide-coated surfaces and is behavior familiar to all workers on photoelectric effects in air.

In theory, the onset of a self-sustaining discharge from a negative point is given by $\gamma \exp \int_a^r \alpha dx = 1$, where $\alpha$ is Townsend's first coefficient, $dx$ is the distance from the center of the point, $r$ its radius, $a$ is the distance from the point at which Townsend's coefficient loses its efficiency in the declining field. $\gamma$ is the chance of emission of electrons from the cathode per ion created in the gap, and could be ascribed to a photo-effect at the surface, or more probably to electron emission by positive ion bombardment. Thus $\gamma$ is critically dependent on condition of the surface and on the work function, which must be less than the ionization potential of the positive ions. From the nature of the current-potential curves for the negative point, it *had* been assumed that the Trichel pulse appearance set the threshold defined above.

For the positive burst pulse threshold, Loeb[6] has shown that $\beta f \exp \int_a^r \alpha dx = 1$, where $f$ is the fraction of the collisions producing ions that give photons capable of ionizing in the gas, and $\beta$ is a geometrical factor dealing with the chance of absorption of such a photon beyond $a$ in the gas. Here onset depends entirely on the properties of the gas and not on the surface. In glow discharges at lower pressures, and in studies of Townsend's coefficients at low pressures, $\gamma$ is of the order of $10^{-2}$ to $10^{-3}$. It has never been evaluated at atmospheric pressure in gases, but owing to back diffusion[7] of emitted electrons alone, it could be smaller at 760 mm. Another factor that cannot be neglected is that $\gamma$ is dependent on the energy of impact of positive ions on the surface. No data below positive ion energies of 20 volts exist. For air, $\beta f$ has the value of about $10^{-4}$. Thus, it is clear that if the Trichel pulse onset evaluates $\gamma$ in air at atmospheric pressure, $\gamma$ must be of the order of $10^{-4}$, while in pure nitrogen and otherwise, it can well be $10^{-2}$ to $10^{-3}$.

However, this is not the worst difficulty the present observations present. They indicate that the Trichel pulse onset determined value of $\gamma$ is independent of the metal or the metal oxide surface. This independence of $\gamma$ for the widely different metals or their oxides was tentatively ascribed to the presence of some common *gaseous* film on all metal surfaces, involving oxygen or nitrogen oxides. That some sort of an adsorbed film does form in the presence of oxygen which is common in properties, and behavior on all surfaces cannot be denied, and copious evidence for the existence of such films exists. However, it is hardly probable that such a film of gas could alter the value of $\gamma$ to the extent of making it completely independent of the metal, as different metals might have films of different $\gamma$.

As long as the work was confined to studies of the corona with point-to-plane geometry, the character of the observed onset curves was such that one could hardly avoid concluding that the Trichel pulse onset was the *actual threshold* of a self-sustaining negative discharge, and that it fixed the value of $\gamma$. The difficulty in a more discriminating decision here, lies in the instrumental difficulties of observing the exact point of Trichel pulse onset relative to the very sharp increase in current which seems to mark the appearance of the pulses. The recent work of C. G. Miller,[8] using concentric cylinder geometry with a fine central wire, and studying gases ranging from pure nitrogen through various mixtures of oxygen and nitrogen to pure oxygen, has disclosed a set of circumstances which now modify the picture. With point-to-plane corona and relatively small points, pre-onset currents remain at about $10^{-13}$ ampere until a potential is reached at which there is an abrupt change of current to about $10^{-7}$ ampere with the appearance of Trichel pulses in the steeply rising portion. With concentric cylinders, depending on the character of the gas, Miller observed that, in general, the rise of current with potential up to the point where Trichel pulses occurred, was far more gradual than with the point. Thus, the jump of current on the appearance of the Trichel pulses was less sudden. The Trichel pulses also first appeared well up on the steep rising portion of the current voltage curve. His results with guard ring cylinders indicated that there was a definite and *clearly defined pre-onset current along the whole wire* in contrast to localized pulses, which in pure oxygen at 760 mm was half as large as the total current when the Trichel pulses appeared. In other mixtures, while the rise was more abrupt, it was less abrupt than any observed with point-to-plane corona. This indicates that *well below the threshold of Trichel pulse onset, the ordinary field intensified gaseous discharge has been replaced by a self-sustaining Townsend discharge* of low current value along the whole length of the wire. Since this Townsend discharge is below an intensity where a glow is visible, and since it develops in a

[6] L. B. Loeb, Phys. Rev. **73**, 798 (1948).
[7] L. B. Loeb, *Fundamental Processes of Electrical Discharge in Gases* (John Wiley & Sons, Inc., New York 1939), pp. 315 and 402.
[8] C. G. Miller, Ph.D. thesis, University of California (1948); Phys. Rev. **75**, 1460 (1949).

current range which makes it inaccessible to oscilloscopic observation, the transition from the intermittent Townsend discharge to the steady self-sustaining Townsend discharge evaluating $\gamma$ cannot be indicated even with Miller's set-up. It is, however, considerably below Trichel pulse onset in potential as judged from the magnitudes of Miller's currents. Thus, the onset of a self-sustaining discharge, which defines the value of $\gamma$, lies at potentials far below Trichel pulse onset. Hence, the onset and thus, $\gamma$ may quite definitely be a function of electrode material as well as of gas coating. Until the true onset can be observed, this point will not be decided. *Trichel pulse onset, therefore, does not represent a threshold defining a* $\gamma$. With a point-to-plane geometry, the Townsend discharge current, owing to the limited area of the point, would be so small that its contribution to the current would hardly be noted. Thus, the abrupt increase of current near Trichel pulse onset was naturally but mistakenly assumed to represent the threshold for self-sustaining discharge.

This conclusion permits immediate interpretation of the present observations. The onset of the self-sustaining Townsend discharge, fixed by the true $\gamma$ of the surface, leads to a type of glow discharge whose value is limited by the relatively low value of $\gamma$. The currents in this limited discharge increase in a regular and consistent fashion with increase in potential, since the first Townsend coefficient is increasing irrespective of any small increase in $\gamma$. It appears, however, that with the low $\gamma$ from the gas-coated surface, the secondary electron emission is so low that the positive ion space charge created by the secondary electrons coming from the point is not sufficient to initiate the autocatalytic build-up of ionization described by Loeb.[1] However, as the potential is raised, the surface of the point or wire is being bombarded by greater densities of positive ions and possibly with positive ions of increasing energy. At higher positive ion energies, such bombardment leads to sputtering of the point material; here it leads to removal of the gas and/or oxide films. There is, then, at each current, an equilibrium process set up in which there is competition between the removal of the gas film by positive ion bombardment and the tendency of the gas film to re-form by the diffusion of oxygen atoms and molecules to the point. Eventually a current density must be reached at which the surface of the metal becomes denuded of its gas and perhaps even of its oxide layers, i.e., the point is cleaned up or "conditioned." At this juncture the work function of the surface is very much decreased, and the value of $\gamma$ may increase by an order of magnitude, i.e., from $10^{-3}$ to $10^{-2}$.

The value of the threshold current density for clean-up will depend primarily on the character of the gas, the percentage of active molecules, and the energy of the incoming positive ions. Therefore, the higher the field, the more concentrated the discharge is by the geometry, or the lower the concentration of oxygen, the lower will be the current density at which the film will be removed.

With the film suddenly dispersed at a given potential, the value of $\gamma$ is such that the potential is far above that for the threshold value of any sort of a discharge of this character, i.e., the point is highly overvolted. Thus there will be a secondary liberation of electrons from the overvolted surface which will produce the autocatalytic increase in current which was described by Loeb.[1] This leads to a burst of current which can either go to arc discharge as in pure nitrogen with no limiting resistor, or which can, in a gas like air, create so many electrons in an explosive fashion that negative ion space charges choke off the discharge producing the Trichel pulse. It is seen that this mechanism completely accounts for the Trichel pulse onset. Through the differentiation of Townsend discharge onset and Trichel pulse onset, this picture disassociates the latter from the evaluation of $\gamma$ and makes possible a variation of $\gamma$ with point material. The mechanism proposed accounts for higher pre-onset currents with other geometries. It accounts for the relative lowering of Trichel pulse onset with point radius through the increased current densities and higher tip fields. It accounts for the aging effects with Pt and other surfaces, and for the more gradual increase of pre-onset current with potential before Trichel pulse onset in pure $O_2$ where the $O_2$ reforms the surface layer.[8] In pure $N_2$, and $N_2$ with small amounts of $O_2$, it explains the sudden current increases and change of discharge form from a diffuse to a concentrated discharge when points are initially subjected to constant current for some time at the same potential.[9] It finally indicates the need for much more work on the location of the true Townsend threshold and the evaluation and study of $\gamma$.

[9] G. L. Weissler, Phys. Rev. **63**, 96 (1943).

# Magnetic Properties of Iron Compacts in Relation to Sintering Temperature*

ROBERT STEINITZ
*American Electro Metal Corporation, Yonkers, New York*
(Received January 3, 1949)

Previous experiments have indicated that the magnetic permeability of sintered iron compacts is to a large extent determined by the final density of the compact. The permeability of five different iron powders has proved to be independent of the origin of the powder, if the compacts were sintered at the same temperature. The results were compared with the theory of Polder and Van Santen and agreed very well with their predicted curve, if the residual pores were assumed to be flat disks. It could be expected that at higher sintering temperatures, the disk-shaped pores would change into a more equiaxed or spherical shape.

Experiments to check this theory have been made, and the results are in good agreement. Compared for identical densities, the permeability of different iron powders is appreciably higher if sintering is done at 1250–1350°C and for 24 hours, instead of at 1150°C for one hour. Different powders, however, now show quite different permeabilities, which could be explained by a different inclination of the powder to form spheroidal pores.

It is shown that coining after high temperature sintering brings the permeability values back to the curve of the disk-shaped flat pores. The effect of the increase in permeability at higher temperature cannot, therefore, be due to a purification process, but must be attributed to a change in the pore shape.

BECAUSE of the limitations of commercial furnaces, iron parts for magnetic applications are usually sintered at a temperature between 1000° and 1150°C. Furthermore, it is possible to hold dimensions to close tolerances at this sintering temperature, as shrinkage is only insignificant. It is well known, however, that, at only 100°C higher temperature, dimensional changes and densification already become much more pronounced. It can be expected that this increased sintering activity at higher temperatures manifests itself not only in an increased and accelerated shrinkage, but also has an influence on the physical properties of the compact. This should be especially true for properties which depend on the shape and distribution of the pores. The modern theory of sintering postulates that pores change to a spherical or almost spherical shape before they disappear completely. A comparison of properties obtained after different sintering temperatures should indicate whether such a spheroidization of pores actually takes place, provided the comparison is made for identical densities. This would eliminate the influence of the amount of porosity or the pore volume, and make the pore shape the determining variable.

The maximum magnetic permeability of an iron compact is a variable, very well suited to follow this sintering behavior. Polder and Van Santen[1] have published a theory on the permeability of porous compacts with special reference to the influence of the shape of the pores. It was shown in a previous paper[2] that iron compacts, sintered at around 1150°C for about 1 hr., have a relationship between maximum permeability and density corresponding very closely to the one predicted by the Dutch authors for flat holes. Such a flat disk-like pore shape is to be expected in compressed parts, if sintering temperature and time are not high enough to change it. However, results published elsewhere indicated that an increase in sintering temperature results in a considerable improvement of the permeability, though no systematic studies of this effect have been published.

The purpose of this paper is to investigate the permeability-density relationship for higher sintering temperatures. It is, therefore, a continuation of work previously published,[2] but the present results can still be considered only as preliminary, due to a lack of sufficient data. Further work in this line is in progress. The results obtained, however, indicate satisfactory agreement with theory.

TABLE I. Characteristics of iron powders (reduced powders).

| | No. 1 | No. 2 | No. 3 | No. 4 |
|---|---|---|---|---|
| Apparent density g/cc | 2.76 | 2.44 | 2.38 | 2.7 |
| Carbon content, % | 0.12 | 0.11 | 0.15 | 0.21 |
| Loss in hydrogen, % | 0.73 | 0.43 | 0.40 | 0.55 |
| Screen analysis | | | | |
| +100 | 0 | 0 | 6.5 | 0 |
| −100 +150 | 17.5 | 27.0 | 19.0 | 7.5 |
| −150 +200 | 34.5 | 25.0 | 16.5 | 23.5 |
| −200 +250 | 15.0 | 12.0 | 9.5 | 15.5 |
| −250 +325 | 13.0 | 13.0 | 11.0 | 16.5 |
| −325 | 20.0 | 23.0 | 37.5 | 37.0 |

## MATERIALS AND PROCEDURES

The raw materials used were four reduced iron powders from different manufacturers. One powder, No. 3, was subjected by the manufacturer to a special process which resulted in almost perfectly spherical particles. The characteristics of the powders are given in Table I. Experiments with electrolytic powders, which were included in the previous investigation,[2] are not yet concluded, so that the results are not reported

* This paper was presented at the International Powder Metallurgy Conference at Graz, Austria, July 12–17, 1948.

[1] D. Polder and J. H. Van Santen, Physica 12, 257 (1946).
[2] R. Steinitz, Trans. A.I.M.E., Tech. Pap. No. 2335 (February, 1948).

here. The powders are all quite similar in particle size distribution and other characteristics.

In order to obtain different densities at identical sintering temperatures, the powders were pressed with various pressures, ranging from 10 to 50 t.s.i. A ring with an average diameter of about 2 in. was directly pressed, so that no machining was necessary afterwards.

Rings proved to be the only type of sample which gave reproducible results in magnetic measurements. These rings were sintered in hydrogen for 24 hr. at 1250°C and 1350°C, respectively, and cooled in the cooling chamber of the furnace.

After sintering, the density of the sample was determined by weighing it in air and water, after the sample was coated with a lacquer to prevent water from entering the pores. The rings were then wound with two windings, a magnetizing winding of about 120 turns and a measuring secondary winding of 25 turns. This resulted, with a current of 5 amp., in a magnetizing field of about 50 oersteds, which was considered sufficient. All measurements were done with a ballistic galvanometer.

## RESULTS

The results of the magnetic measurements of all samples are given in Table II, together with sintering temperature and time and resulting density. The results for a sintering temperature of 1150°C were either taken from the previous investigation[2] or from new measurements, where new powders were used. In order to get well reproducible results, it proved necessary to demagnetize with special care. Due to the effect of eddy currents, especially in dense materials, the maximum amperage of the demagnetizing a.c. must be higher than the maximum d.c. used, and should be decreased continuously to a very low value before being switched off. The measurement of the maximum permeability on the initial magnetization curve is very sensitive to any remaining magnetism. But even if all precautions are taken, results still fluctuate and stray values with considerable deviations occur. Another source of error is the density determination, as well as the determination of length and cross section of the ring sample. As the sintering temperature has a pronounced effect on the permeability, a uniform temperature in the sintering furnace is required.

## DISCUSSION OF RESULTS

In Fig. 1, the maximum permeability for all four powders after sintering for 1 hr. at 1150°C is reproduced, as function of the density. (The density is given in percent, 7.85 g/cc being taken as the value for 100 percent dense iron.) The values for powders Nos. 1, 2, and 3 are all well on the curve which was taken from the previous publication. The one value for powder No. 4, however, is somewhat higher; this is an exception to the observations on all other powders, the permeability values for which are all well represented by this curve. A reason for the behavior of this powder is not yet known; however, it might be noted that the shrinkage in this case is already appreciably higher than with the "normal" powders. It is possible that the sintering temperature of 1150°C is already quite high for powder No. 4.

TABLE II. Magnetic properties of compacts made from reduced powders.

| Powder | Sintering | Density % | Residual magnetism $B_r$ | Coercive force $H_c$ | Maximum permeability $\mu_{max}$ | Induction $B_{40}$ |
|---|---|---|---|---|---|---|
| No. 1 | 1150°C 1 hr. | 93 | 11,900 | 3.5 | 1590 | 14,300 |
| | | 89 | 11,600 | 3.5 | 1600 | 13,400 |
| | | 81 | 7000 | 4.5 | 700 | 9800 |
| | | 78 | 6400 | 4.3 | 700 | 8800 |
| | | 75 | 6500 | 4.5 | 730 | 8200 |
| | 1250°C 24 hr. | 90 | 10,200 | 1.9 | 2500 | 12,800 |
| | | 88.5 | 11,400 | 1.8 | 3100 | 13,800 |
| | | 83.5 | 9300 | 2 | 2100 | 11,300 |
| | | 80.5 | 8300 | 2.1 | 1900 | 10,600 |
| | | 79 | 8500 | 2 | 2100 | 10,200 |
| | 1350°C 24 hr. | 92 | 10,800 | 1.4 | 3000 | 13,500 |
| | | 82 | 9200 | 1.8 | 2000 | 11,000 |
| | | 77 | 7300 | 1.8 | 1800 | 9400 |
| No. 2 | 1150°C 1 hr. | 90.5 | 9600 | 2.8 | 1470 | 13,000 |
| | | 74 | 4900 | 3.3 | 680 | 7000 |
| | 1250°C 24 hr. | 89 | 8100 | 3.3 | 1350 | 13,000 |
| | | 74.5 | 5500 | 3.6 | 850 | 8700 |
| No. 3 | 1150°C 1 hr. | 71 | 5500 | 3.7 | 750 | 7800 |
| | 1250°C 24 hr. | 88.5 | 9500 | 2.1 | 2000 | 13,000 |
| | | 86.5 | 8500 | 3 | 2000 | 12,400 |
| | | 79.5 | 7500 | 3.2 | 1500 | 10,100 |
| | 1350°C 24 hr. | 93.5 | 10,400 | 1 | 2600 | 14,900 |
| | | 90 | 8300 | 2 | 2100 | 12,500 |
| | | 73.5 | 6150 | 2 | 1500 | 8800 |
| No. 4 | 1150°C 1 hr. | 80 | 7500 | 2.5 | 1400 | 11,000 |
| | 1250°C 24 hr. | 93 | 12,250 | 1.6 | 3300 | 15,000 |
| | | 86 | 11,000 | 1.5 | 3100 | 13,200 |
| | | 79.5 | 9300 | 1.9 | 2350 | 11,500 |
| | 1350°C 24 hr. | 95 | 10,000 | 1.6 | 3200 | 15,500 |
| | | 86 | 10,100 | 1.5 | 3150 | 13,000 |
| | | 79.5 | 7800 | 1.8 | 2500 | 10,400 |

Figure 2 gives the values of the maximum permeability after sintering for 24 hr. at 1250°C. The points for the 4 powders are no longer on one curve; on the contrary, they differ appreciably, powder No. 4 having about 3 times the permeability of powder No. 2. Despite considerable fluctuation, the general trend of the curve for each powder can be easily recognized. For powder No. 1, this curve gives permeability values about 2½ times higher than after 1150°C sintering, which is a very considerable improvement.

The sintering at 1250°C was extended for 24 hr.,

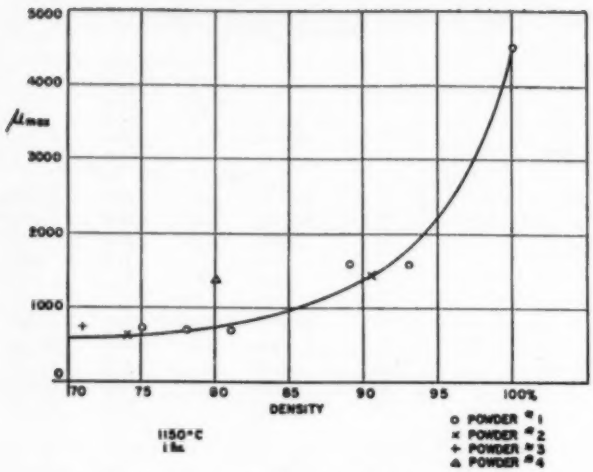


FIG. 1. Permeability-density relationship for sintering for 1 hr. at 1150°C.

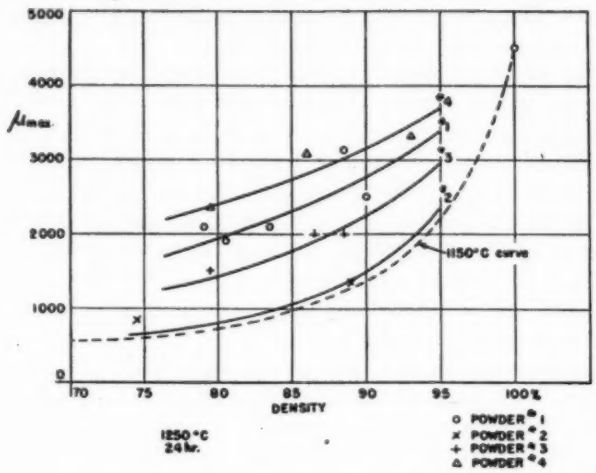


FIG. 2. Permeability-density relationship for sintering for 24 hr. at 1250°C.

while, at 1150°C, the sintering time was only 1 hr. This was done to make more pronounced any effect the higher temperature might have. However, a consequence of this procedure is that there is little difference between 1250°C and 1350°C sintering. Values for both temperatures are practically on the same curve. Some kind of an equilibrium appears to have been reached. The change of the permeability with time for the same sintering temperatures is still being investigated.

The induction value for a magnetizing field of 40 oersteds, $B_{40}$, was taken as a measure for the saturation. It is interesting to note that these induction values are not influenced by the sintering temperature. They all lie well within the expected fluctuations on one curve, if plotted against density, as is true for the residual magnetism $B_r$. No trend with temperature is discernible.

It was claimed that the improvement in permeability is due to the purification by loss of carbon at the higher sintering temperatures. That this is not the essential effect can be seen from the following: A ring was sintered at about 1350°C, then coined and annealed at 760°C. This resulted in a density of 87.5 percent and a maximum permeability of 1530, which is close to the curve for 1150°C sintering. (See reference 2, Table II, Specimen No. 12.) After the high sintering temperature, a much higher permeability would be expected for this powder. Any purification would, of course, still be effective after coining, but it is believed that the coining process flattens the pores again and thereby removes any spheroidization effect of the high sintering temperature.

## SUMMARY AND CONCLUSIONS

Permeabilities of iron compacts prepared in the form of rings were measured after sintering at 1250° and 1350°C for 24 hr., and compared with values of samples sintered at lower temperatures. A considerable increase of the permeability for identical densities was found for most of the powders, but values for different powders no longer lay on the same curve.

An explanation for the increase in permeability is given by the theory of Polder and Van Santen,[1] if it is assumed that the pores change their shape from flat disks to approximately spheroids at the high sintering temperature. An attempt to follow this change in pore shape under the microscope was not successful. The pores are probably too small and too easily changed in the polishing procedure. The difference in the behavior of the permeability of the various powders cannot yet be completely explained. It is probably not due to a difference in chemical composition, but to a greater or lesser sintering activity of the powder which is also shown by an increased shrinkage. Measurements of relative shrinkage and behavior of mechanical as well as magnetic properties at different sintering temperatures and times, always compared for equal densities, will help to explain these phenomena and show whether the sintering process and the change in pore shape can be followed by these experiments. The improvement of the magnetic properties by the higher sintering temperature will also be of commercial importance.

# The Vaporization of the Cathode in the Electric Arc

RAGNAR HOLM
*Stackpole Carbon Company, St. Marys, Pennsylvania*
(Received March 7, 1949)

Calculations are carried out which show that, because of the smallness of the cathode spot, the cathode is not able to dissipate the heat generated with heavy currents by conduction. Therefore, a much higher vaporization of cathode material per coulomb occurs for heavy currents than for low currents.

## I. INTRODUCTION

THE following computations provide an explanation for the fact that the vaporization per coulomb from the cathode of an arc is greater at heavy currents than at low currents. The treatment is based on observations and on the assumptions that the temperature in the cathode spot is of the order of the boiling point of the cathode material. All of the observations on the arc, which have been used for the calculations, are of an approximate nature, and do not motivate exact calculations. Therefore, several simplifications have been used. One of the aims of this paper is to call the attention to a problem where new and more accurate observations would make possible important conclusions.

For information about some details, references are made to the book, *Electric Contacts*,[1] where further literature is listed.

## II. THE ENERGY BALANCE AT THE CATHODE SPOT

We neglect the relatively small radiation from the cathode spot, and assume that the total energy produced in the cathode drop region is transferred to the cathode.

The *gain of energy*, $W$, in the spot is given by the product of the current, $I$, and a voltage which is the difference between the cathode drop, $V_c$, and the heat of vaporization of the electrons given by the work function, $\varphi$. Hence approximately

$$\left.\begin{array}{l}\text{for metals } W=(V_c-\varphi)I=10I\\ \text{and}\\ \text{for carbon } W=15I\end{array}\right\}. \tag{1}$$

For simplicity, the calculations are carried out for metals only in the following derivation. We regard the cathode spot as a circle with the radius $a$ and designate the current density by $j$. Thus

$$I=\pi a^2 j. \tag{2}$$

It is important to notice that $j$ is fairly uniform across the cathode spot because of the uniformity of the temperature. Consequently, $j$ in Eq. (2) may be regarded as a constant for the concerned metal and current intensity.

[1] Ragnar Holm, *Electric Contacts* (H. Geber, Stockholm, 1946). This book is denoted E.C.H. in the following.

The *losses of energy* per second are the heat $W_c$ conducted away from the spot into the cathode, and the heat $W_v$ used for vaporization of cathode material.

The heat $W_c$ is equal to the temperature $\vartheta$ of the cathode spot times the thermal conductance away from it. We regard the cathode as a semi-infinite body. The conductance[2] is then $4a\lambda$, where $\lambda$ is the heat conductivity and

$$W_c=4a\lambda\vartheta. \tag{3}$$

The latent heat of vaporization may be $H$ joule/cm$^3$ and the volume evaporated per second $v$ cm$^3$/sec. Thus

$$W_v=Hv. \tag{4}$$

The energy balance requires that

$$W=W_c+W_v$$

or

$$10I=10\pi a^2 j=4a\lambda\vartheta+Hv. \tag{5}$$

We see that the radius $a$ of the spot enters with different powers in two terms of Eq. (5). This fact is the basis for important conclusions. It means that, with increasing current $I$ and radius $a$, the term $10\pi a^2 j$ increases quicker than the term $4a\lambda\vartheta$. Finally, at large currents, the term $4a\lambda\vartheta$ becomes negligible, and all the heat is used for vaporization.

We shall follow the calculations through in some detail. First we introduce two variables:

$\delta=v/I$, the evaporated volume per coulomb removed from the cathode by the arc, and
$x=$ the fraction of the total energy which is used for vaporization;

thus

$$xW=W_v=Hv$$

or

$$\delta=10x/H. \tag{6}$$

We see that the volume $v$ may be replaced by $I\delta$ in

TABLE I.

| Materials | | Cu | W | C | |
|---|---|---|---|---|---|
| $\vartheta$ | = | 2300 | 5000 | 5000 | °C |
| $\lambda$ | = | 3.8 | 1.5 | 0.04 | watt/cm deg |
| $H$ | = | 5 | 9.2 | 21 | $10^4$ joule/cm$^3$ |

[2] See reference 1, p. 23.

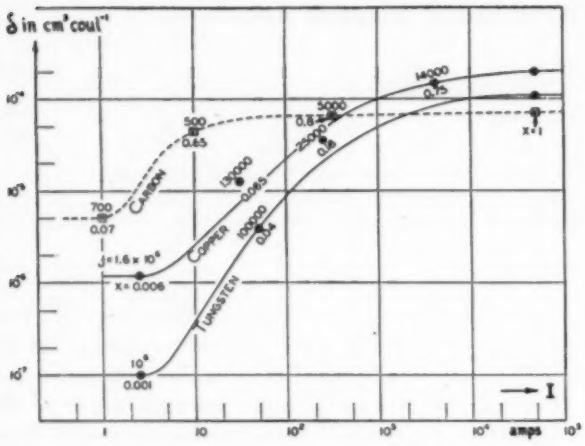


FIG. 1. Volume $\delta$ of the cathode material vaporized per coulomb through the arc plotted against the current intensity. The number written above a plotted point gives the current density $j$ amp./cm²; the number below the point gives $x$, the fraction of the energy developed in the cathode drop region, which is available for vaporization.

Eq. (5). Then Eqs. (2), (5) and (6) have the variables

$$a, j, x, \delta, \lambda, H, \vartheta, I.$$

The equations determine three of these variables if the others are known. Values of $I$ may be assumed. The values of $\vartheta$, $\lambda$ and $H$ which have been used are given in Table I.

As for $a$, $x$, $\delta$, and $j$ we must regard one of them to be known. For some points plotted in Fig. 1 $\delta$ has been measured.[3] Thus, for these points the radius $a$ and $j$ and $x$ can be calculated. For the largest currents, for instance, $I=5\times10^4$ amps. we regard $4a\lambda\vartheta$ as negligible and put $x=1$. Then the equations are able to determine $\delta$, but $j$ and $a$ remain undetermined, since under this condition Eqs. (2) and (6) do not tell more than Eq. (5).

In the case of carbon $j$ is known,[4] and $a$, $\delta$ and $x$ can be calculated.

### III. NUMERICAL RESULTS

The results of the calculations are illustrated in Fig. 1. The number written above a plotted point gives $j$; the number under the point gives $x$.

All curves begin at small currents with a horizontal branch and end at very large currents in another horizontal branch. It may be noticed that a horizontal line here indicates that the volume evaporated per second is proportional to the current $I$.

The lower horizontal branch is an interesting singularity. It represents measurements cited in E.C.H.[5] Here only a small fraction of the energy which is carried by the ions to the cathode is used for vaporization. The metal atoms obtain the energy for their escape directly from the impinging positive ions. I would say, they are kicked out, and this effect is independent of the temperature of the cathode spot. In fact, it seems[6] that the fraction of the energy of the ions, which is used for the vaporization, or call it the disintegration or sputtering of the cathode, is nearly the same at the cool cathode spot of a glow discharge and the hot spot of an arc, at a small current. A. von Hippel[7] has discussed this independence of the temperature as far as the glow discharge is concerned, but it is valid also for arcs.

The kicking out of atoms will probably give the clue to the explanation of the high speed jets of metal vapor that have been observed at arc cathodes by Tanberg and Haynes.[8] The atoms in these jets have taken over more energy from the ions than is needed for the escape.

The horizontal branch at large currents is a consequence of the condition that $x=1$. Here the vaporization is primarily a normal thermal one. Since the term $4a\lambda\vartheta$ in Eq. (5) is negligible, the conductivity $\lambda$ does not affect this branch. The metal property which is decisive is $H$, the heat of vaporization, and it is not able to cause very great differences for different materials.

This high vaporization from the cathode is, of course, a nuisance for the constructors of switches, but it is inevitable if our assumptions are right. The vaporization from the anode and splashing of molten metal are other disturbing phenomena which shall not be discussed here.

Finally, attention may be called to the high values of the current density $j$ at small currents, which have been indicated in Fig. 1. They are consequences of the assumptions made. Recent measurements[9] indicate the existence of such current densities in transient arcs.

[3] Some of these measurements are new and some are taken from E.C.H. Table (57.09) after conversion from cathode loss to cathode evaporation according to E.C.H. Eq. (57.03).

[4] W. Finkelnburg, *Hochstromkohlebogen* (Springer-Verlag, Heidelberg, 1948), p. 18; or "The high current carbon arc," FIAT Final Report 1052, PB-81644, Office of Technical Services, Department of Commerce, Washington, D. C., 1948, p. 11.

[5] See E.C.H. paragraph 57, particularly Tables (57.13) and (57.14) and Eq. (57.03), also Appendices I and IV.

[6] See reference 1, paragraph 57 E.

[7] The concerned papers by A. v. Hippel are cited and criticized by N. D. Morgulis, M. P. Bernardiner, and A. M. Patiocha. Physik. Zeits. Sowjetunion **9**, 302 (1936); see also E.C.H. p. 316.

[8] See J. R. Haynes, Phys. Rev. **73**, 891 (1948).

[9] See H. Roth, Dissertation, Dresden 1939; L. H. Germer and collaborators, Lectures given at the gas discharge conference, Brookhaven National Laboratory, Oct. 28, 1948; J. D. Cobine and C. J. Gallacher, Phys. Rev. **74**, 1524 (1948).

# A Study of the e.m.f. Method*

C. T. Tai
*Cruft Laboratory, Harvard University, Cambridge, Massachusetts*
(Received September 15, 1948)

The present work contains an investigation of the e.m.f. method to determine the input impedance of a thin biconical antenna as first suggested by Schelkunoff. By considering the various components of the currents flowing on the antenna, it has been shown that the sinusoidal part is, indeed, the dominating component among all these currents. An investigation of the total current at the end of the antenna, where the lateral surface of the cone and the spherical cap meet, shows that the current is not identically zero at that point, but vanishes as the reciprocal of the square of the characteristic impedance of the cone. The two approximate expressions of the complex conjugate power used in the e.m.f. method are derived, as well as the exact expression. Computation of some integrals occurring in the formulation is also treated in detail to show the various approximations involved.

THE e.m.f. method was first suggested by Brillouin[1,2] to compute the power radiated by an antenna, or by a system of antennas. It is an alternative to the standard Poynting vector method. The difference between the two is that the latter deals with the integration of the Poynting vector over a big surface enclosing the system, while the former is based on direct integration of the Poynting vector along the conductor. As was shown by Bechmann,[3] both methods are equivalent and one can be converted into the other by means of Poynting's theorem. In this investigation, we will consider the applicability of the e.m.f. method to the determination not only of the power radiated by antenna but also its input impedance. This application was first studied by Schelkunoff[4] in his theory of biconical antennas, where he derived an expression for the effective load admittance of a thin biconical antenna based upon two apparently different methods, of which one is the e.m.f. method. A recent investigation[5] of the same problem shows that Schelkunoff's expression can be derived using a more rigorous method based upon Smith's formulation of the problem,[6] and involving fewer assumptions than Schelkunoff's original work. We have thus verified a result obtained by the e.m.f. method. What remains to be done, therefore and the object of the present work, is to give some interpretations and justifications of the various assumptions involved in that method.

* The research reported in this document was made possible through support extended Cruft Laboratory, Harvard University, jointly by the Navy Department (ONR) and the Signal Corps, U. S. Army, under Contract N5ori-76, T. O. 1.

[1] L. Brillouin, "Sur l'origine de la resistance de rayonnement," Radioelectricite **3**, 147 (1922).

[2] A. A. Pistolkors, "The Radiation Resistance of Beam Antennas," Proc. I.R.E. **17**, 562 (1929).

[3] R. Bechmann, "On the Calculation of Radiation Resistance of Antennas and Antenna Combinations," Proc. I.R.E. **19**, 1471 (1931).

[4] S. A. Schelkunoff, "Theory of Antennas of Arbitrary Size and Shape," Proc. I.R.E. **29**, 493 (1941).

[5] C. T. Tai, "On the Theory of Biconical Antennas," Cruft Laboratory Technical Report No. 47, June 10, 1948; J. App. Phys. **19**, 1155 (1948).

[6] P. D. P. Smith, "On the Conical Dipole of Wide Angle," J. App. Phys. **19**, 11 (1948).

## FIELDS ASSOCIATED WITH A BICONICAL ANTENNA

Without going into the details of the derivation, we will simply write down most of the equations derived previously.[4–6] The principal wave existing in region $I$ of Fig. 1 is given by

$$E_\theta=\frac{-jI_0}{4\pi r\sin\theta}\Big[(1+KY_t)e^{-j\beta(r-l)}+(1-KY_t)e^{j\beta(r-l)}\Big], \quad (1)^{**}$$

$$H_\phi=\frac{-jI_0}{4\pi r\sin\theta}\Big[(1+KY_t)e^{-j\beta(r-l)}-(1-KY_t)e^{j\beta(r-l)}\Big], \quad (2)$$

where $I_0$, as will be shown later, is the amplitude of the principal current on the antenna; $Y_t$ is the effective admittance acting as a load upon the finite conical transmission line formed by the lateral surfaces of the cones; and $K$ is the characteristic impedance of the cones, being given by

$$K=z_0/\pi\ (\log\cot\theta_1/2); \quad (3)$$

where

$$z_0=(\mu_0/\epsilon_0)^{\frac{1}{2}}\doteq 120\pi \text{ ohms}. \quad (4)$$

The complementary waves existing in region $I$ are

$$E_r=\frac{-j}{\omega\epsilon_0 r^2}\sum_n\frac{a_n}{2\pi}\frac{S_n(\beta r)}{S_n(\beta l)}L_n(\theta), \quad (5)$$

$$E_\theta=\frac{-jz_0}{r}\sum_n\frac{a_n}{2\pi}\cdot\frac{1}{n(n+1)}\cdot\frac{S_n'(\beta r)}{S_n(\beta l)}\cdot\frac{\partial L_n(\theta)}{\partial\theta}, \quad (6)$$

$$H_\phi=\frac{-1}{r}\sum_n\frac{a_n}{2\pi}\cdot\frac{1}{n(n+1)}\cdot\frac{S_n(\beta r)}{S_n(\beta l)}\cdot\frac{\partial L_n(\theta)}{\partial\theta}, \quad (7)$$

** The arbitrary constant $V$ occurring in Eqs. (1) and (2) of reference 5 has been replaced by $-jI_0/4\pi$ in this paper, where $I_0$ is a new constant having the dimension of current. The change is necessary in order to make our notation of the principal current the same as that of Schelkunoff.

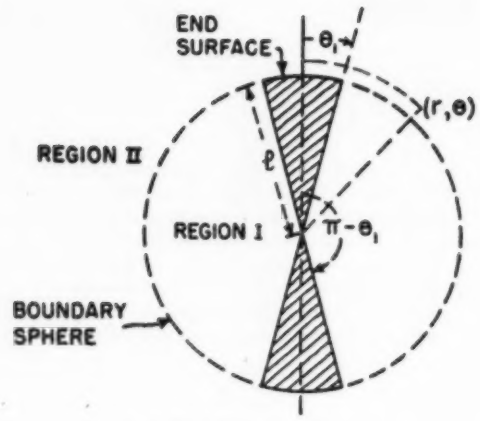


FIG. 1. The biconical antenna.

where

$$S_n(\beta r)=(\beta r)^{\frac{1}{2}}J_{n+\frac{1}{2}}(\beta r) \tag{8}$$

$$S_n'(\beta r)=[dS_n(\beta r)/d(\beta r)]. \tag{9}$$

The complementary waves existing in region $II$ are given by

$$E_r=\frac{-j}{\omega\epsilon_0 r^2}\sum_{k=1,3\cdots}\frac{b_k}{2\pi}\frac{R_k(\beta r)}{R_k(\beta l)}P_k(\theta), \tag{10}$$

$$E_\theta=\frac{-jz_0}{r}\sum_{k=1,3\cdots}\frac{b_k}{2\pi}\cdot\frac{1}{k(k+1)}\frac{R_k'(\beta r)}{R_k(\beta l)}\frac{\partial P_k(\theta)}{\partial\theta}, \tag{11}$$

$$H_\phi=\frac{-1}{r}\sum_{k=1,3\cdots}\frac{b_k}{2\pi}\frac{1}{k(k+1)}\frac{R_k(\beta r)}{R_k(\beta l)}\frac{\partial P_k(\theta)}{\partial\theta}, \tag{12}$$

where

$$R_k(\beta r)=(\beta r)^{\frac{1}{2}}H_{k+\frac{1}{2}}^{(2)}(\beta r), \tag{13}$$

$$R_k'(\beta r)=[dR_k(\beta r)/d(\beta r)]. \tag{14}$$

It was shown by Schelkunoff[4] that the problem of determining the input impedance is essentially that of determining $Y_t$. Once this quantity is found the input impedance of the antenna is given by

$$Z_{in}=\left[\frac{KY_t\sin\beta l-j\cos\beta l}{\sin\beta l-jKY_t\cos\beta l}\right]K. \tag{15}$$

By matching the $H_\phi$- and $E_\theta$-components of the interior and exterior fields across the boundary sphere and at the end-surfaces, the following set of equations can be derived:[5,6]

$$Y_t=\frac{jz_0}{\pi K^2}\sum_{k=1,3\cdots}\left(\frac{b_k}{I_0}\right)\frac{1}{k(k+1)}P_k(\mu_1), \tag{16}$$

$$-jP_r(\mu_1)=\left(\frac{b_r}{I_0}\right)W_r+r(r+1)P_r(\mu_1)$$
$$\times\sum_{k=1,3\cdots}\sum_n\left(\frac{b_k}{I_0}\right)\frac{P_k(\mu_1)s_n}{(r,n)(k,n)},\quad r=1,3\cdots, \tag{17}$$

where

$$\mu_1=\cos\theta_1, \tag{18}$$

$$s_n=j\frac{2n+1}{n(n+1)}(1-\mu_1^2)\frac{dn}{d\mu_1}\frac{S_n'(\beta l)}{S_n(\beta l)}, \tag{19}$$

$$(r,n)=r^2+r-n^2-n, \tag{20}$$

$$W_r=\frac{j}{2r+1}\frac{R_r'(\beta l)}{R_r(\beta l)}. \tag{21}$$

Equations (16) and (17) contain the formal solution for $Y_t$ for biconical antennas of any angle. For small-angle cones, it has been shown[5] that Eq. (17) yields a rather simple solution for the coefficient $(b_r/I_0)$, which is given by

$$\frac{b_r}{I\epsilon_0}=-j\frac{\pi}{2}(2r+1)\beta l J_{r+\frac{1}{2}}(\beta l)H_{r+\frac{1}{2}}^{(2)}(\beta l); \tag{22}$$

hence the corresponding value of $Y_t$ is

$$Y_t=\frac{z_0\beta l}{2K^2}\sum_{k=1,3\cdots}\frac{2k+1}{k(k+1)}J_{k+\frac{1}{2}}(\beta l)H_{k+\frac{1}{2}}^{(2)}(\beta l). \tag{23}$$

For convenience, we may write

$$K^2Y_t=Z_a, \tag{24}$$

with

$$Z_a=R_a+jX_a$$
$$=\frac{z_0}{2}\beta l\sum_{k=1,3\cdots}\frac{2k+1}{k(k+1)}J_{k+\frac{1}{2}}(\beta l)H_{k+\frac{1}{2}}^{(2)}(\beta l). \tag{25}$$

The above formulation forms the basis of the ensuring discussion of the e.m.f. method.

## GENERAL EXPRESSIONS OF THE VOLTAGE AND THE CURRENT

It is obvious that in order to apply the e.m.f. method to finding the power radiated from the antenna or to determining its input impedance, we need the expressions of both the voltage and the current acting upon the antenna. As shown by Schelkunoff,[4] if we define the voltage as the line integral of the $E_\theta$-component along a meridian, then the interior complementary waves contribute nothing to the voltage thus defined, while the principal wave gives

$$V(r)=KI_0[KY_t\sin\beta(l-r)-j\cos\beta(l-r)]. \tag{26}$$

The current flowing in the antenna can best be described by separating it into several parts. On the lateral surface of the antenna, the current associated with the principal wave is given by

$$I(r)=I_0[\sin\beta(l-r)-jKY_t\cos\beta(l-r)]. \tag{27}$$

Equation (27) is obtained by putting $\theta=\theta_1$ in Eq. (2) and then multiplying the resultant expression of $H_\phi$ by

$2\pi r \sin\theta_1$. The current associated with the interior complementary wave is obtained by using Eq. (7) instead of (2). That gives

$$I_c(r)=-\sum_n \frac{a_n}{n(n+1)}\frac{S_n(\beta r)}{S_n(\beta l)}\sin\theta_1\frac{\partial L_n(\theta_1)}{\partial\theta_1}$$

$$=\sum_n \frac{a_n}{n(n+1)}\frac{S_n(\beta r)}{S_n(\beta l)}(1-\mu_1^2)\frac{\partial L_n(\mu_1)}{\partial\mu_1}. \quad (28)$$

In order to investigate the relative magnitudes of the currents defined by Eqs. (27) and (28), it is necessary to transform Eq. (28) into a form containing the coefficients $b_k$ instead of $a_n$. It was shown by Smith[6] that as a result of matching the interior and exterior fields across the boundary sphere, the following relation exists:

$$a_n I_{nn'}=\sum_{k=1,3\cdots} b_k I_{nk}, \quad (29)$$

where

$$I_{nk}=\frac{1-\mu_1^2}{(k,n)}P_k(\mu_1)\frac{\partial L_n(\mu_1)}{\partial\mu_1}, \quad (30)$$

$$I_{nn'}=\frac{1-\mu_1^2}{2n+1}\frac{\partial L_n(\mu_1)}{\partial n}\cdot\frac{\partial L_n(\mu_1)}{\partial\mu_1}. \quad (31)$$

By making use of the relation

$$\frac{\partial L_n(\mu_1)}{\partial n}\cdot\frac{dn}{d\mu_1}+\frac{\partial L_n(\mu_1)}{\partial\mu_1}=0 \quad (32)$$

and Eqs. (29)–(31), there is no difficulty in changing Eq. (28) into the following form:

$$I_c(r)=-\sum_n\sum_{k=1,3\cdots} b_k\frac{2n+1}{n(n+1)}\frac{1}{(k,n)}$$

$$\times(1-\mu_1^2)P_k(\mu_1)\frac{dn}{d\mu_1}\frac{S_n(\beta r)}{S_n(\beta l)},$$

$$=-\sum_n\sum_{k=1,3\cdots}\frac{b_k}{k(k+1)}\left[\frac{1}{n}-\frac{1}{n-k}+\frac{1}{n+1}\right.$$

$$\left.-\frac{1}{n+k+1}\right](1-\mu_1^2)P_k(\mu_1)\frac{dn}{d\mu_1}\frac{S_n(\beta r)}{S_n(\beta l)}. \quad (33)$$

On the spherical end-surfaces of the cones, the current is equal to $2\pi l\sin\theta H_\phi$, where $H_\phi$ is the magnetic field defined in region *II* and given by Eq. (12). Thus, we have

$$I_t(\theta)=-\sum_{k=1,3\cdots} b_k\frac{1}{k(k+1)}\sin\theta\frac{\partial P_k(\theta)}{\partial\theta},$$

$$=\sum_{k=1,3\cdots} b_k\frac{1}{k(k+1)}(1-\mu^2)\frac{\partial P_k(\mu)}{\partial\mu}, \quad (34)$$

where $\mu=\cos\theta$. Equations (27), (33) and (34) are the general expressions of the currents flowing on the lateral surface and on the end-surfaces of a biconical antenna. Their specific values depend upon the amplitudes of the exterior complementary waves $b_k$ which satisfy an infinite set of linear equations as given by Eq. (17).

## CURRENTS ON SMALL-ANGLE CONES

For small-angle cones we have shown that the value of $b_k$ is given by Eq. (22) and that of $Y_t$ by Eqs. (24)–(25). The principal current defined by Eq. (27) can then be written as

$$I(r)=I_0[\sin\beta(l-r)-j(Z_a/K)\cos\beta(l-r)]. \quad (35)$$

For convenience, the following notations will be used:

$$I_1(r)=I_0\sin\beta(l-r) \quad (36)$$

$$I_2(r)=-j(Z_aI_0/K)\cos\beta(l-r). \quad (37)$$

For $I_c(r)$, it has been shown[5] that when $\mu_1\to 1$, the following relations exist:

$$\lim_{\mu_1\to 1}(1-\mu_1^2)\frac{dn}{d\mu_1}=-(n-k)^2$$

$$=-\frac{1}{[\log(2/\theta_1)]^2}=-\frac{z_0^2}{\pi^2K^2}, \quad (38)$$

$$\lim_{\mu_1\to 1}(n-k)=[1/\log(2/\theta_1)]=z_0/\pi K, \quad (39)$$

$$\lim_{\mu_1\to 1}P_k(\mu_1)=1. \quad (40)$$

As a result, $I_c(r)$ can be split into two parts; the first corresponds to the sum of those terms of the double series that contain the factor $1/(n-k)$ for values of $k$ nearly equal to $n$, while the remainder constitute the second part. Denoted respectively by $I_3(r)$ and $I_4(r)$, they are

$$I_3(r)=\frac{-z_0}{\pi K}\sum_{k=1,3\cdots}\frac{b_k}{k(k+1)}\frac{S_k(\beta r)}{S_k(\beta l)}, \quad (41)$$

$$I_4(r)=\frac{z_0^2}{\pi^2K^2}\left[-\sum_{\substack{n=1,3\cdots\\(n\neq k)}}\sum_{k=1,3\cdots}\frac{b_k}{k(k+1)}\frac{1}{(n-k)}\frac{S_n(\beta r)}{S_n(\beta l)}\right.$$

$$+\sum_{n=1,3\cdots}\sum_{k=1,3\cdots}\frac{b_k}{k(k+1)}\left[\frac{1}{n}+\frac{1}{n+1}\right.$$

$$\left.\left.-\frac{1}{n+k+1}\right]\frac{S_n(\beta r)}{S_n(\beta l)}\right]. \quad (42)$$

Substituting the value of $b_k$ as given by Eq. (22) into

(41), we have

$$I_3(r)=j\frac{z_0I_0}{2K}\sum_{k=1,3\cdots}\frac{2k+1}{k(k+1)}$$

$$\times\left[(\beta l)^{\frac{1}{2}}H^{(2)}_{k+\frac{1}{2}}(\beta l)\right]\left[(\beta r)^{\frac{1}{2}}J_{k+\frac{1}{2}}(\beta r)\right]. \quad (43)$$

A similar expression can be derived for $I_4$, which is then proportional to $I_0/K^2$. Thus, if $I_1$ is designated as the zero*th* order distribution of current on the antenna, both $I_2$ and $I_3$ then have a weighting factor that is proportional to $1/K$, while the weighting factor of $I_4$ is proportional to $1/K^2$. At the feeding point $r=0$, both $I_3(0)$ and $I_4(0)$ are identically zero. The input current is therefore contributed by $I_1(0)$ and $I_2(0)$ only. At $r=l$, $I_1(l)$ is zero but $I_2(l)$, $I_3(l)$, and $I_4(l)$ are, in general, nonvanishing. From Eqs. (37), (43) and (25) it is seen that

$$I_2(l)=-j(Z_aI_0/K), \quad (44)$$

and

$$I_3(l)=j(Z_aI_0/K). \quad (45)$$

For $I_4(l)$, by putting $r=l$ in Eq. (42) and summing the series with respect to $n$, the following expression can be

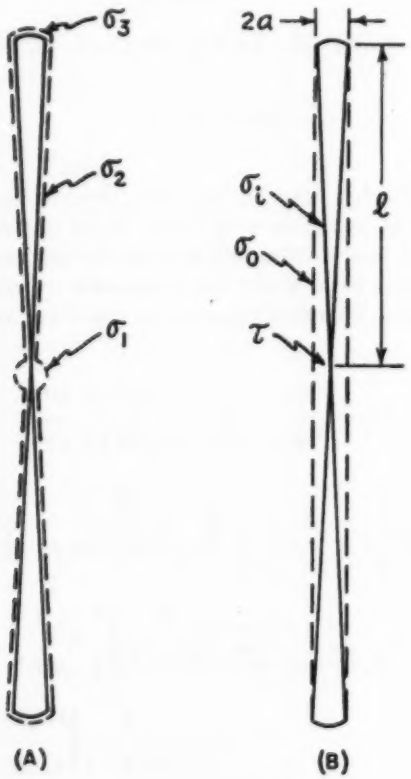


FIG. 2. Small-angle biconical antenna showing the various surfaces.

derived:

$$I_4(l)=\frac{z_0^2}{\pi^2K^2}\sum_{k=1,3\cdots}\frac{b_k}{k(k+1)}[C+\psi(k+1)]=\frac{-j\beta lz_0^2I_0}{2\pi K^2}$$

$$\times\sum_{k=1,3\cdots}\frac{(2k+1)[C+\psi(k+1)]}{k(k+1)}J_{k+\frac{1}{2}}(\beta l)H^{(2)}_{k+\frac{1}{2}}(\beta l), \quad (46)$$

where $C$ is the Euler's constant and $\psi(k+1)$ is the logarithmic derivative of $\Gamma(k+1)$. Since the sum of $I_2(l)$ and $I_3(l)$ is zero, we conclude that the total axial current at $r=l$ vanishes as $1/K^2$, provided the series occurring in Eq. (46) is convergent. The latter statement will be postulated here pending a mathematical proof. In passing we may observe that in Schelkunoff's method of partial matching,[4,5] the vanishing of the total current at the end of the antenna is assumed as one intermediate step in his analysis. The actual picture now acquires a greater clarity from what has been discussed here. Knowing the relative magnitudes of the various currents associated with the principal and complementary waves, we can proceed to a discussion of the parts taken by these currents in the e.m.f. method.

## POWER RELATIONS

The essential idea in applying the e.m.f. method to find the input impedance of a small-angle biconical antenna is based upon a comparison of two different expressions of the complex power fed into the system. These power relations will be discussed under separate headings.

### A. Exact Expression for the Complex Conjugate Power Fed into the Antenna

By means of Eqs. (15), (26) and (27) we find that the complex conjugate power fed into the antenna is given by

$$P=\tfrac{1}{2}V(0)\cdot I^*(0)=\tfrac{1}{2}|I(0)|^2Z_{in}$$

$$=\tfrac{1}{2}|I_0|^2\left[K^2G_t+j\left(\frac{-K\sin 2\beta l}{2}\right.\right.$$

$$\left.\left.-K^2B_t\cos 2\beta l+K^3|Y_t|^2\frac{2\sin 2\beta l}{2}\right)\right], \quad (47)$$

where $G_t$ and $B_t$ are respectively the conductance and susceptance of $Y_t$; i.e.,

$$Y_t=G_t+jB_t. \quad (48)$$

As will be shown later, in making use of the e.m.f. method to find the input impedance of the antenna, instead of this exact expression *it is the approximate expression of the complex conjugate power which will be effective.*

### B. Approximate Expression for the Complex Conjugate Power Fed into the Antenna

By approximate expression for the complex conjugate power we mean that expression of $P$ in which an approximate expression of $I(0)$ is used but a more exact expression of $Z_{in}$ is retained. Thus for $I(0)$ in $\frac{1}{2}I(0)^2Z_{in}$ we use $I_0\sin\beta l$, but neglect the term $-jI_0KY_t\cos\beta l$ in Eq. (27); and for $Z_{in}$ we expand the denominator of Eq. (15) using Taylor's theorem and neglect the high-order terms. This gives

$$Z_{in}\doteq\frac{1}{\sin^2\beta l}\left[K^2G_t+j\left(\frac{-K\sin\beta l}{2}+K^2B_t\right)\right]. \quad (49)$$

Hence,

$$P_{app}=\tfrac{1}{2}|I_0|^2\left[K^2G_t+j\left(\frac{-K\sin\beta l}{2}+K^2B_t\right)\right]. \quad (50)$$

Comparing Eqs. (47) and (50) it will be noted that in addition to the term $(K^2|Y_t|^2\sin2\beta l)/2$, which is negligibly small, the coefficients associated with $B_t$ are quite different for the two cases. This is of considerable consequence in determining $Y_t$ be means of the e.m.f. method, as will be seen later.

### C. Expression for the Complex Conjugate Power Derived from Poynting's Theorem

The foregoing expressions for the complex conjugate power are derived by regarding the antenna simply as a circuit element. On the other hand, an equivalent expression for the complex conjugate power can be derived by applying the Poynting theorem and integrating the total power along the surface of the conductor. Referring to Fig. 2(A), we observe that the total complex conjugate power radiated from the antenna is given by

$$P=\tfrac{1}{2}\int_{\sigma_1,\sigma_2,\sigma_3}E\cdot H^*d\sigma \quad (51)$$

where $\sigma_1$ is an infinitesimal spherical surface surrounding the generator, $\sigma_2$ is the lateral surface of the cones and $\sigma_3$ their top surface. If we use the exact expressions of $E$ and $H$ to compute $P$, then, since the tangential component of $E$ is zero along $\sigma_2$ and $\sigma_3$, the only contribution is due to the integration along $\sigma_1$. It can be shown without any difficulty that by using (1) and (2) an expression for $P$ is derived that is identically the same as Eq. (47). There is, however, another approach to the problem that enables us to find an approximate expression for the complex conjugate power based upon Poynting's theorem.

Referring back to Eq. (51), let us decompose the integral into several parts and express them in different forms. Thus if we denote

$$P_t=\tfrac{1}{2}\int_{\sigma_1,\sigma_2}E\cdot H^*d\sigma=\tfrac{1}{2}\int_{\sigma_1,\sigma_2}E_r\cdot H_\phi{}^*d\sigma \quad (52)$$

and

$$P_t=\tfrac{1}{2}\int_{\sigma_3}E\cdot H^*d\sigma=\tfrac{1}{2}\int_{\sigma_3}E_\theta\cdot H_\phi{}^*d\sigma, \quad (53)$$

then

$$P=P_t+P_t. \quad (54)$$

For $P_t$, one has

$$P_t=\tfrac{1}{2}\int_{\sigma_1,\sigma_2}E_r\cdot H_\phi{}^*d\sigma=\tfrac{1}{2}\int_{-l}^{l}2\pi r\sin\theta_1E_rH_\phi{}^*dr$$

$$=\tfrac{1}{2}\int_{-l}^{l}I_z{}^*E_rdr, \quad (55)$$

where $I_z$ is the total current flowing along the lateral surface of the cone. For $P_t$, one has

$$P_t=\frac{1}{2}\int_{\sigma_3}E_\theta H_\phi{}^*d\sigma=\int_0^{\theta_1}2\pi l^2\sin\theta E_\theta H_\phi{}^*d\theta$$

$$=l\int_0^{\theta_1}I_\theta{}^*E_\theta d\theta \quad (56)$$

where $I_\theta$ is the total radial current flowing on the top surface of the cone. It will be recalled that for small-angle biconical antennas, $I_z$ can be separated into four parts given by Eqs. (36), (37), (41) and (42), while $I_\theta$ is given by Eq. (34). If we think of each of these currents as a source of radiation, then there is an electromagnetic field associated with each of these currents. As a result, we may write

$$E_r=\sum_{n=1}^{5}E_{nr} \quad (57)$$

and

$$E_\theta=\sum_{n=1}^{5}E_{n\theta}, \quad (58)$$

where, for example, $E_{1r}$ is the $r$-component of the $E$-field, expressed in a spherical coordinate system, produced by $I_1$; and $E_{5r}$, $E_{5\theta}$ are the components due to $I_\theta$, which will henceforth be denoted by $I_5$. Substituting these formal expressions of $I_z$, $I_\theta$, $E_r$ and $E_\theta$ into Eqs. (55) and (56), we obtain

$$P_t=\frac{1}{2}\int_{-l}^{l}\left(\sum_{k=1}^{4}I_k{}^*\sum_{n=1}^{5}E_{nr}\right)dr \quad (59)$$

and

$$P_t=l\int_0^{\theta_1}\left(I_5{}^*\sum_{n=1}^{5}E_{n\theta}\right)d\theta. \quad (60)$$

Of all the cross-products involved in Eqs. (59) and (60), because of the different orders of magnitude of the various currents, it is obvious that the most significant part is due to the integration of $I_1{}^*E_{1r}$. The remaining terms, though contributing only negligibly to the value of the total complex conjugate power radiated from the system, nevertheless are important

in the sense that by their inclusion the boundary condition is automatically satisfied.

The evaluation of the integral

$$P_{app}=\frac{1}{2}\int_{-l}^{l} I_1^* E_{1r}\,dr \tag{61}$$

was briefly described by Schelkunoff both in his paper[4] and in his book,[7] but since the computation definitely involves some approximations, it is perhaps worthy of a more detailed derivation here. Moreover, it will be seen that the steps adopted here do not precisely follow those outlined by Schelkunoff. It is well known[8] that the field produced by a sinusoidal distribution of current

$$I_1=I_0\sin\beta(l-|z|) \tag{62}$$

is given by

$$E_{1z}=\frac{jI_0z_0}{4\pi}\left[\frac{-e^{-j\beta R_1}}{R_1}-\frac{e^{-j\beta R_2}}{R_2}+\frac{2}{R_0}\cos\beta l e^{-j\beta R_0}\right] \tag{63}$$

$$E_{1\rho}=\frac{jI_0z_0}{4\pi\rho}\left[\frac{z-l}{R_1}e^{-j\beta R_1}+\frac{z+l}{R_2}e^{-j\beta R_2}-\frac{2z\cos\beta l}{R_0}e^{-j\beta R_0}\right] \tag{64}$$

$$H_\phi=\frac{jI_0}{4\pi\rho}\left[e^{-j\beta R_1}+e^{-j\beta R_2}-2\cos\beta l e^{-j\beta R_0}\right] \tag{65}$$

where

$$\left.\begin{aligned} R_1&=((l-z)^2+\rho^2)^{\frac{1}{2}}\\ R_2&=((l+z)^2+\rho^2)^{\frac{1}{2}}\\ R_0&=(z^2+\rho^2)^{\frac{1}{2}}\end{aligned}\right\} \tag{66}$$

and $\rho$ is the distance measured from the point of observation to the $z$-axis. According to Schelkunoff,[7] one can use $E_{1z}+E_{1\rho}\sin\theta_1$ for $E_{1r}$ in Eq. (61) to perform the integration. The resulting integral, however, is somewhat complicated. A simpler method is to perform the integration along a cylindrical surface as shown in Fig. 2(B), and then to make the necessary correction for a conical surface by means of Poynting's theorem. Denoting the lateral surface of the cone by $\sigma_i$, the surface of the open cylinder by $\sigma_0$, and the volume in between these two surfaces by $\tau$, we find[9] that

$$\frac{1}{2}\int_{\sigma_i} E\cdot H^*d\sigma=\frac{1}{2}\int_{\sigma_0} E\cdot H^*d\sigma+j\frac{\omega}{2}\int_{\tau}(\mu H^2-\epsilon E^2)d\tau \tag{67}$$

or

$$\frac{1}{2}\int_{-l}^{l} I_1^*E_{1r}\,dr=\frac{1}{2}\int_{-l}^{l} I_1^*E_{1z}\,dz+j\frac{\omega}{2}\int_{(a/l)z}^{a}\int_{-l}^{l} 2\pi\rho(\mu H^2-\epsilon E^2)d\rho dz. \tag{68}$$

The two integrals at the right-hand side of Eq. (68) can be evaluated as shown in the appendix, giving

$$\frac{1}{2}\int_{-l}^{l} I_1^*E_{1r}\,dr=\tfrac{1}{2}|I_0|^2(R+jX) \tag{69}$$

where

$$R=z_0/4\pi[2(C+\log 2\beta l-Ci2\beta l)+(Si4\beta l-2Si2Bl)\sin 2\beta l+(C+\log\beta l-2Ci2\beta l+Ci4\beta l)\cos 2\beta l] \tag{70}$$

$$X=z_0/4\pi[2Si2\beta l-Si4\beta l\cos 2\beta l+(-C-\log\beta l-2\log 2\beta l+2\log\beta a+Ci4\beta l)\sin 2\beta l]. \tag{71}$$

For small-angle cones, as treated here, the characteristic impedance of the cones is given by

$$K\doteq(z_0/\pi)\log(2/\theta_1)=(z_0/\pi)\log(2\beta l/\beta a). \tag{72}$$

Thus, $X$ can be written as

$$X=(-K\sin 2\beta l/2)+z_0/4\pi[2Si2\beta l-Si4\beta l\cos 2\beta l+(-C-\log\beta l+Ci4\beta l)\sin 2\beta l]. \tag{73}$$

Now we have two *approximate* expressions of the power fed into a thin biconical antenna; one is given formally by Eq. (50), while the other has been derived from Poynting's theorem and is given by Eqs. (69), (70) and (73). If they are equivalent we should obtain

$$K^2G_t=z_0/4\pi[2(C+\log 2\beta l-Ci2\beta l)+(Si4\beta l-2Si2\beta l)\sin 2\beta l+(C+\log\beta l-2Ci2\beta l+Ci4\beta l)\cos 2\beta l] \tag{74}$$

and

$$K^2B_t=z_0/4\pi[2Si2\beta l-Si4\beta l\cos 2\beta l-(C+\log\beta l-Ci4\beta l)\sin 2\beta l]. \tag{75}$$

As shown by Rice,[10] the expression of $K^2Y_t$ thus obtained is identically equal to the real and imaginary part of $Z_a$ given by Eq. (25), which was obtained more rigorously without resorting to the various approximations encountered in the e.m.f. method. It is to be noticed that although we use the approximate expressions of the complex conjugate power to determine the value of $K^2Y_t$, the actual power fed to the system is given by Eq. (47) instead of by Eq. (50) or Eq. (69). On the other hand, if we identify Eq. (47) with Eq. (69) the value of $K^2B_t$ thus determined would not be the same as before and it would not be the correct one, even though the value of $K^2G_t$ is unaffected. This is an indication of the critical nature of the e.m.f. method: Unless it is handled with extreme care an error may easily result.

## APPENDIX

### Evaluation of the Integral $\frac{1}{2}\int_{-l}^{l} I_1^*E_{1r}\,dr$

[7] S. A. Schelkunoff, *Electromagnetic Waves* (D. Van Nostrand Company, Inc., New York, 1943), p. 453.

[8] See reference 7, p. 371.

[9] R. W. P. King, *Electromagnetic Engineering* (McGraw-Hill Book-Company, Inc., New York, 1945), Vol. I, p. 221.

[10] S. O. Rice, "Sum of Series of the Form $\Sigma_0^\infty A_n J_{n+\alpha}(z)J_{n+\beta}(z)$," Phil. Mag. XXXV, 686 (1944).

The first integral at the right-hand side of Eq. (68) is known*** to have the following value, for $a^2 \ll l^2$.

$$\frac{1}{2}\int_{-l}^{l} I_1^* E_{1z} = \tfrac{1}{2}|I_0|^2(R+jX_c) \qquad \text{(A-1)}$$

where

$$R=\frac{z_0}{4\pi}[2(C+\log 2\beta l - Ci2\beta l)+(Si4\beta l - 2Si2\beta l)\sin 2\beta l + (C+\log\beta l - 2Ci2\beta l + Ci4\beta l)\cos 2\beta l] \qquad \text{(A-2)}$$

$$X_c=\frac{z_0}{4\pi}[2Si2\beta l+(2Si2\beta l - Si4\beta l)\cos 2\beta l + (C-\log\beta l+2\log\beta a - 2Ci2\beta l + Ci4\beta l)\sin 2\beta l]. \qquad \text{(A-3)}$$

To evaluate the second integral of Eq. (68), we use the expressions of $E$ and $H$ given by Eqs. (63)–(65). It can be shown that

$$\mu H^2 - \epsilon E^2 = \frac{\mu I_0^2}{8\pi^2\rho^2}\left[\left(1-\frac{R_0^2-l^2}{R_1R_2}\right)\cos(R_2-R_1) - 2\left(1-\frac{R_0^2-zl}{R_0R_1}\right)\cos\beta l\cos\beta(R_1-R_0) - 2\left(1-\frac{R_0^2+zl}{R_0R_1}\right)\cos\beta l\cos\beta(R_2-R_0)\right]. \qquad \text{(A-4)}$$

When $\rho \ll z$, $(l-z)$, $(l+z)$, the following approximations are permissible:

$$1-\frac{R_0^2-l^2}{R_1R_2}\doteq 2, \quad R_2-R_1\doteq 2z$$

$$1-\frac{R_0^2-zl}{R_0R_1}\doteq 2, \quad R_1-R_0\doteq l-2z$$

$$1-\frac{R_0^2+zl}{R_0R_2}\doteq 0, \quad R_2-R_0\doteq l.$$

Then Eq. (A-4) reduces to

$$\mu H^2-\epsilon E^2\doteq(-\mu|I_0|^2/4\pi^2\rho^2)\cos 2\beta(l-z). \qquad \text{(A-5)}$$

Substituting Eq. (A-5) into the second integral of Eq. (68) and performing the integration, one obtains

$$j\frac{\omega}{2}\int_{(a/l)z}^{a}\int_{-l}^{l} 2\pi\rho(\mu H^2-\epsilon E^2)\,d\rho\,dz = \tfrac{1}{2}|I_0|^2\Delta X \qquad \text{(A-6)}$$

where

$$\Delta X=-j\frac{z_0}{4\pi}[2Si2\beta l\cos 2\beta l+2(C-Ci2\beta l+\log 2\beta l)\sin 2\beta l]. \qquad \text{(A-7)}$$

Thus, we have

$$\frac{1}{2}\int_{-l}^{l} I_1^* E_{1r}\,dr = \tfrac{1}{2}|I_0|^2(R+jX) \qquad \text{(A-8)}$$

where $R$ is given by Eq. (A-2) and

$$X=X_c+\Delta X=\frac{z_0}{4\pi}[2Si2\beta l - Si4\beta l\cos 2\beta l + (-C-\log\beta l-2\log 2\beta l+2\log\beta a+Ci4\beta l)\sin 2\beta l]. \qquad \text{(A-9}$$

*** S. A. Schelkunoff, *Electromagnetic Waves* (D. Van Nostrand Company, Inc., New York, 1943), pp. 372–373.

## BOOK REVIEW

### Atomic Theory for Students of Metallurgy, Institute of Metals Monograph No. 3

By William Hume-Rothery. Pp. 282, Figs. 124, 7s.6d. The Institute of Metals, London, 1946.

This monograph presents a summary of the mathematical theories of the electronic structure of atoms, molecules, crystals and alloy systems in language understandable to the advanced, non-mathematically minded student. The accomplishments of wave mechanics as applied to the fundamental questions that explain the structure of solids, or that show promise of ultimately doing so are presented with remarkable clarity. The author has an unusual gift for this kind of writing, and has a breadth of first hand information on the nature of alloys that comes from years of research directed to understanding the fundamental principles governing the structure of alloys. Although not a mathematical physicist himself, he has studied this field extensively and has been guided past many pitfalls in drawing conclusions from the published theories by discussions with Professor Mott and Dr. Frolich, as he says in a footnote.

The book outlines many of the metallurgical problems, the successes and some of the inadequacies of theory in a way most helpful to the metallurgist and the average graduate student in the sciences; it does not attempt to derive any of the principles, but merely to explain, interpret, apply, and test them. More attention is given to the successes than to the failures of theories, as is understandable.

There is no extensive overlapping of the contents with the contents of Hume-Rothery's "Structure of Metals and Alloys," another monograph of the same series (revised 1947). The subjects treated include the Bohr theory, the wave equations and their interpretation, potential boundaries in wave mechanics, the electronic structure of the atom and assemblies of atoms, the interactions between atoms in molecules and crystals, Fermi-Dirac statistics and the electron gas, the Bruillion zone theory, and the application of the theories to the metals, particularly copper, silver, gold, the hexagonal metals, and the elements of the first long period. The fields of elasticity, anelasticity, plastic deformation, and preferred orientations are not covered.

C. S. Barrett
*Institute for the Study of Metals, The University of Chicago*

# Letters to the Editor

## Ionospheric Absorption

R. N. Bracewell and K. Weekes
*Cavendish Laboratory, Cambridge, England*
April 11, 1949

DOUBT has been cast, by your correspondent L. G. McCracken,[1] on the formula given by Best and Ratcliffe[2] for the absorption of a radio wave which penetrates an ionized region of the Chapman type. He states that an approximation has been made, as a result of which the formula is correct only when the sun is near the horizon. He states that when the sun is overhead, the formula is wrong by a factor of approximately one quarter. The formula is, however, well-established and is in agreement with the results obtained in a different way by Appleton.[3]

The formula given by Best and Ratcliffe, under the assumption that the wave frequency is much greater than the collision frequency, comes out in terms of the error function as follows:

$$-\tfrac{1}{2}\log\rho=\frac{H}{2mc}\frac{4\pi e^2\nu_0}{(p+|p_L|)^2}\left(\frac{B}{\alpha}\right)^{\frac{1}{2}}\left(\frac{2}{C\sec\chi}\right)^{\frac{1}{2}}\times\left[u\exp(-u^2)-\int\exp(-u^2)du\right]_{(\frac{1}{2}C\sec\chi)^{\frac{1}{2}}}^{0}. \quad (1)$$

McCracken obtains the following formula under the same assumption:

$$-\tfrac{1}{2}\log\rho=\frac{H}{2mc}\frac{4\pi e^2\nu_0}{(p+|p_L|)^2}\left(\frac{B}{\alpha}\right)^{\frac{1}{2}}(\tfrac{1}{2},\tfrac{1}{2}C\sec\chi)!$$

where $(x, y)!$ is the incomplete factorial function. Thus where the first formula has

$$\int_0^{(\frac{1}{2}C\sec\chi)^{\frac{1}{2}}}\exp(-u^2)du-(\tfrac{1}{2}C\sec\chi)^{\frac{1}{2}}\exp(-\tfrac{1}{2}C\sec\chi) \quad (2)$$

the second has

$$(\tfrac{1}{2},\tfrac{1}{2}C\sec\chi)!$$

Now it is a property of the incomplete factorial function that

$$(\tfrac{1}{2},a)!=\tfrac{1}{2}(-\tfrac{1}{2},a)!-a^{\frac{1}{2}}e^{-a}.$$

Further, since

$$(n^{-1}-1,x^n)!=n\int_0^x\exp(-t^n)dt,$$

it follows that

$$\tfrac{1}{2}(-\tfrac{1}{2},a)!=\int_0^{a^{\frac{1}{2}}}\exp(-t^2)dt.$$

Hence:

$$(\tfrac{1}{2},a)!=\int_0^{a^{\frac{1}{2}}}\exp(-t^2)dt-a^{\frac{1}{2}}e^{-a}.$$

The expression (3) is therefore equal to (2) if we identify $a$ with $\frac{1}{2}C\sec\chi$, and McCraken's formula is identical with that of Best and Ratcliffe.

The numerical discrepancy mentioned above arose because McCracken erroneously took $(\frac{1}{2},\frac{1}{2}C\sec\chi)!$ equal to $(\frac{1}{2},\frac{1}{2})!$ when $\chi=0$, thus implying that $C=1$. Now $(\frac{1}{2},\frac{1}{2})!$ is smaller than the correct value by a factor 0.199 which gives rise to the discrepancy stated to be approximately one quarter.

Best and Ratcliffe do not, after setting up their integral, "immediately go to an approximation," but correctly perform the integration, obtaining an expression in the form (1), and this form is simpler than that of McCracken. To evaluate their expression they discard the second term in (2) and take the upper limit of the error integral to be infinite. The change introduced is small—about $e^{-78,400}$.

These adjustments were numerical, but have the following physical significance. The variable $u$ in (1) comes from the substitution $u=(\frac{1}{2}C\sec\chi)^{\frac{1}{2}}\exp(-\frac{1}{2}h/H)$, where $h$ is the height above the ground, $H$ is the scale height of the atmosphere, $\chi$ is the sun's zenith distance, and $C\sec\chi=\exp(h_m/H)$, where $h_m$ is the height of maximum ionization density. The limits of integration, $(\frac{1}{2}C\sec\chi)^{\frac{1}{2}}$ and 0, in (1), signify that the integral is taken from the ground level to infinite height. If the integration is taken from negatively infinite height upwards to infinity, as throughout Appleton's treatment, the absorption varies as $(\cos\chi)^{\frac{1}{2}}$. But in the general case of integration upwards from some finite level, the rising and falling of the region, as the sun's zenith distance changes, leads to a departure from the $(\cos\chi)^{\frac{1}{2}}$ law. In the practical case of the ionospheric layers, the height of maximum ionization for region $E$ is about twelve scale heights, and $C\sec\chi$ thus has the large value $e^{12}$. The result is almost identical with that obtained by assuming that the integral extends over all values of height, and this is what Best and Ratcliffe get in replacing (2) by $\frac{1}{2}\pi^{\frac{1}{2}}$.

Appleton's treatment makes the same assumption, in that he takes the integral over all values of height. He also evaluates the integral numerically for any relative value of wave frequency and collision frequency.

It may be noticed that the case of $C=1$, to which McCracken's conclusion applies, corresponds to a layer whose maximum ionization density is at ground level when the sun is overhead, and it is to be expected that such a case would lead to a severe departure from the $(\cos\chi)^{\frac{1}{2}}$ relation.

[1] L. G. McCracken, J. App. Phys. **20**, 229 (1949).
[2] J. E. Best and J. A. Ratcliffe, Proc. Phys. Soc. London **50**, 233 (1938).
[3] E. V. Appleton, Proc. Roy. Soc. **A162**, 451 (1937).

## Errata: On Theoretical Signal-to-Noise Ratios in F-M Receivers: A Comparison with Amplitude Modulation

[J. App. Phys. **20**, 334 (1949)]

David Middleton
*Cruft Laboratory, Harvard University, Cambridge, Massachusetts*

Equation (4.2): $\dot{\theta}$ should read $\dot{\theta}^2$.

Equation (5.20): the left number is $(s/n)^2_{III-IIa}$. Omit the word "where" immediately preceding Eq. (II.2).

Equation (II.7): $\Omega^2$ should read $\pi^2$. In the sentence immediately preceding Eq. (II.8), $\phi$ and $\psi$ should be replaced by $\chi$ and $\nu$, respectively.

Equation (II.10): replace $\Psi$ by $\nu$. Write $\langle B_1(R)\dot{\theta}\rangle$ in the sentence following Eq. (II.10).

Equation (II.13): Replace $\phi$ by $\chi$, and in the equation immediately following write $\nu$ for $\psi$; $\sin(\phi-\psi)$ should read $\sin(\chi-\nu)$. In the first sentence following Eq. (II.31) we should have $2\dot{\phi}^2$ *vs.* $\rho^2/a_0^2$, instead of $2\dot{\phi}^2$ *vs.* $\rho^2/a_0^1$.

## Erratum: The Thermal Behavior of Evaporated Layers in Vacuum Devices

[J. App. Phys. **20**, 8 (1949)]

H. A. Stahl
*Laboratory for Material Physics, Telefunken Tube Factory, Berlin, Germany*

THE author has pointed to an uncompleted reference. On page 9, the following should be added to footnote 2: I. Mathes, Zeits. f. tech. Physik **22**, 232 (1941).